

nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

RIBBON DEVELOPMENT

Spontaneous self-assembly of
graphene sheets to form
multi-layer ribbons

PAGE 271

BOOKS

SUMMER READING

*Holiday picks from
Nature's reviewers*

PAGE 228

PUBLIC HEALTH

A NEW ERA FOR HIV

*South Africa's experience of
lifetime antiretroviral therapy*

PAGE 214

POLICY MAKING

SEEKING GREEN DEVELOPMENT

*Recipe for a sustainable
global bioeconomy*

PAGE 221

NATURE.COM/NATURE

14 July 2016 £10

Vol 535, No 7611



THIS WEEK

EDITORIALS

PAIN This will hurt us more than it does you **p.200**

WORLD VIEW The new cheats who game impact metrics **p.201**



MONKEY NUTS Old stone anvils used to shell cashews **p.203**

Closing the label door

The US Senate has agreed a curious compromise on how to identify GM foods. Researchers and policymakers must now tackle more pressing issues with the technology.

Of all the debates over genetically modified (GM) crops, arguments on the need for labels to identify GM food might seem one of the more trivial. From a scientific point of view, by the time a product has reached the shelves, the various tests and standards have long assured its safety. Safety seems the only rational reason for shoppers to reject a food, and therefore the only need for a label.

The problem with that attitude is that it helps to explain why the lack of labelling of GM foods has created so much controversy. If consumers feel that they are being denied a choice, then they tend to object. Hiding information about ingredients has made consumers wonder why it was hidden. It has created an atmosphere that has fostered conspiracy theories, not a deeper understanding of the issues at hand.

Last week, the US Senate passed a bill that will finally create federal standards for GM labels. Widely expected to be passed by the House of Representatives, the bill is clearly a political compromise. Like many good solutions to complex problems, it leaves both sides in the debate feeling hard done by. The law gives federal regulators responsibility to develop mandatory labelling standards, but does not require labels to be printed directly on the products. Instead, consumers can be directed to a website, for example.

It is a solution that could create fairly obscure labels and fully satisfies neither of the two vigorous and vocal sides in the debate. Those are the activists, who argue that consumers should have ready access to information about their food, and the industry lobbyists, who argue that such labels would unfairly taint GM foods — products that a panel convened by the US National Academies of Sciences, Engineering, and Medicine reported, again, in May are safe to eat.

The curious decision reflects the pressure of the atmosphere in which it was forged. Members of the US Senate agricultural committee have been scrambling to find a palatable national standard for months, under intense pressure from industry. The clock was ticking: on 1 July, Vermont became the first US state to enact a law governing such labels, and food manufacturers faced an emerging and confusing regulatory patchwork as other states followed suit.

Now, at least, those who are motivated will be able to find the information they seek. And because federal law trumps state regulations, the new system seems more workable and sensible. The state initiatives seem unusually broad given that the country grows so many GM crops. Whereas some countries provide exemptions for ingredients that are present in trace amounts, or for foods in which the product of genetic engineering — for example, the protein responsible for tolerance to an herbicide — is no longer present, Vermont's law provided no such distinction.

The Senate compromise also promises to address the patchwork problem. The Vermont law, for example, could not supersede the US

Department of Agriculture's authority over certain meat products. As a result, a frozen cheese pizza could require a label if it contains oil made from transgenic soya beans, whereas the same pizza with added pepperoni might not. Federal legislation would do away with such artificial and bewildering distinctions.

But perhaps more importantly, quenching the labelling debate could open the door to discussions about more pressing matters.

“Quenching the labelling debate could open the door to discussions about more pressing matters.”

Vast resources — both time and money — have been poured into the labelling debate. There is an opportunity now to redirect those resources.

In June, the US Environmental Protection Agency's inspector general determined that the agency was not doing enough to cope with the rise of insects resistant to the pesticide produced by some GM crops. Superweeds that are resistant to the herbicides used on certain GM crops are also plaguing farms. And sophisticated gene-editing technologies are helping to bring a new breed of engineered crops to market, yet regulators are still grappling with how to handle them.

Each of these issues is steeped in complex science, and researchers should seize every opportunity to inform — and encourage — discussions around them. The battle over labels has been bruising, and many researchers are hesitant to enter the fray surrounding GM crops. But without their input, the discussion is unlikely to progress. ■

Lend me your ears

A study of how people perceive music shows that jarring chords are a cultural contrivance.

Writing about music has been compared to dancing about architecture, but bear with us.

Santa Maria is a village in western Bolivia without running water or electricity, and so remote that it can be reached only by canoeing up a tributary of the Amazon. It is home to the Tsimane' people, who detect no difference between consonant and dissonant sounds — the relationships between notes that make, for example, 'Eleanor Rigby' by The Beatles sound so sad.

Dissonant chords are the unstable isotopes of Western music; they sound tense and want to revert to more stable forms. The way that composers create and resolve this tension usually invokes different moods in the listener. But not in the Tsimane'.

As researchers describe in a paper this week, when they tested

the musical discrimination of the 'Tsimane' villagers, the listeners experienced consonant and dissonant intervals as equally pleasant (J. H. McDermott *et al.* *Nature* <http://dx.doi.org/10.1038/nature18635>; 2016). This is not a deficiency of affect, because the villagers can distinguish cheerful sounds (laughter) from less cheerful ones (gasps). They also recognize physically unpleasant sonic 'roughness' — the beating sensation when two tones close in frequency are played at once.

The reason for the villagers' inability to distinguish what others would call pleasant sounds from unpleasant ones might be, in large measure, one of culture. The 'Tsimane' do have music, but it is purely one of melody rather than harmony. They play or sing in single lines, and do not adhere to Western scales. This seems odd to those immersed in the European musical tradition, with its clear differences between pleasant and disagreeable harmonies.

The differences are so clear, in fact, that we are inclined to think of them as innate. The mathematics behind the music seems to back this up. Consonant intervals, such as an octave, perfect fourth or perfect fifth, are integral ratios of harmonics — 2:1, 4:3 and 3:2, respectively. A reliably dissonant interval such as the augmented fourth, or tritone, has an irrational ratio of $\sqrt{2}$:1. Consonance and dissonance seem to be written into the fabric of the Universe. But the 'Tsimane' results show that these structures are a human interpretation, and one that seems to be learned by experience.

The tale of the 'Tsimane' should remind us that Western music was not always as richly polyphonic as it is now. In medieval times, music was as melodic as that of the 'Tsimane'. Chords were unknown, and so were modern musical scales. There were just eight notes, corresponding to the white notes on a keyboard. The earliest keyboard instruments had no black keys, and indeed no such thing as a musical key. Instead, there were 'modes', each determined by the unequal spacing of intervals, depending on which note you started from.

But then the Devil arrived, in jumps of three whole tones, in particular between F and B. This was the tritone, so obnoxious that

ecclesiastical authorities described it as *diabolus in musica* ('the Devil in music') and banned it. Choristers presented with singing a tritone preferred to flatten the B, making a much more agreeable perfect fourth. Keyboard technology caught up by inserting the first black key, a B flat. The other black keys followed in time, and modal music evolved into the system of keys that we have today, followed rapidly by that most daring of innovations — polyphony.

"The 'Tsimane' of Bolivia know nothing of Bernstein, let alone Birtwistle."

It is fair to say that the entire edifice of Western music has been built on the tension between consonance and dissonance. The music of Beethoven and Queen's 'Bohemian Rhapsody' take the listener on journeys that make sense only within that framework. Composers Harrison Birtwistle and Pierre Boulez travel routes that redefine the meaning of dissonance and (it must be acknowledged) thrill smaller audiences. Most readers of *Nature*, we hope, can resonate with the heartache and absolution in the song 'Maria' from Leonard Bernstein's *West Side Story*, in which Tony sings the name of his innamorata — using a tritone that immediately resolves into a perfect fifth.

The 'Tsimane' of Bolivia know nothing of Bernstein, let alone Birtwistle. Even when their traditional tunes were recorded, shifted in pitch and harmonized to make polyphonic arrangements and create consonance and dissonance, the listeners could not tell the difference between the two. One hopes that their patience wasn't tried too sorely by outsiders playing fast and loose with their heritage (there are those of us who still bear the scars of hearing Bach murdered by The Beach Boys).

But the key finding, the resolution, the crescendo, the cadenza, the Tierce de Picardie — one is tempted to say — is that the 'Tsimane' do not find the tritone any more or less pleasant than any other interval. The Devil has not reached that part of Bolivia, it seems, and the tunes of the 'Tsimane' might be such as those played in Eden. ■

Fifty shades of pain

The push to find reliable ways to measure pain is proving harder than generating it.

Science has produced such a bewildering array of tools and techniques to cause gentle pain that to list them all can seem like describing a torture chamber in Toytown. To study the body's responses, people are prodded with fingers, pricked with needles and pressed with ice. Toes are squeezed and ear lobes pinched. Muscles can be poked with sticks and zapped with electricity. Mustard oil is spread on the skin and capsaicin injected beneath it. Laser pulses offer a double hit: an initial prick followed by a burning sensation.

When properly performed, these human experimental pain models help researchers to understand both the mechanisms of pain and the effectiveness of new compounds that could help to relieve it. The translational bridge from animal experiments to human trials is built on the backs of countless volunteers who sign up for a little lab-based agony. (Special thanks indeed must go to the anonymous 18 brave souls who had "two series of rectal balloon distensions performed on two separate days" to help to study "cortical processing of visceral sensations and pain" (D. Lelic *et al.* *Neurogastroenterol. Motil.* 27, 832–840; 2015).)

Similar studies check on the pain caused by fully inflating a balloon inside other internal organs. Although, as a review of these pain models noted in 2012, it is (perhaps counter-intuitively) more difficult to find people who are willing to take such balloons through the mouth to stretch the oesophagus: "Difficulties in tolerating balloon

distension commonly results in poor recruitment rates as well as the potential for esophageal perforation" (K. S. Reddy *et al.* *J. Res. Med. Sci.* 17, 587–595; 2012).

When it comes to assessing, measuring and reducing pain, the science toolbox is less well stocked. We have thankfully moved on from the earnest 1950s debates about how the pain tolerance of patients was linked to eye colour — discussions that were themselves coloured by racism. But there is much about pain that we still do not realize, and important knowledge remains beyond the reach of even the best-placed balloon.

Some of what we do know is presented this week in an Outlook supplement (www.nature.com/pain). A series of articles describes the physical, neurological and psychological factors that seem to contribute, and offers a snapshot of current thinking on the best forms of relief.

Science and medicine no longer use a person's ethnicity and religion to mark how well they will tolerate the pain of a procedure, but equally, researchers have not yet found a reliable way to measure pain tolerance. The search for quantifiable ways to compare painful sensations, and to diagnose pain in those who are unable to communicate it, mirrors the effort in psychiatric research to find useful biomarkers for mental-health disorders.

For pain, expression of inflammatory mediators in the blood and the presence of metabolites in saliva could be biological guides to a person's distress. So, too, could brain scans that reveal the neural signature of chronic pain. However, as *Nature* pointed out last year (*Nature* 518, 456; 2015), such techniques must be introduced with care, not least because they could be used by insurance companies and others to demand 'proof' of pain as a way to overrule reported personal experience.

Science has already developed some weird and wonderful ways to deliberately cause pain. It should be wary that it does not inadvertently create some more. ■



Watch out for cheats in citation game

The focus on impact of published research has created new opportunities for misconduct and fraudsters, says Mario Biagioli.

When scientists misbehave, the culture of ‘publish or perish’ is often blamed. Some researchers cut corners, massage data and images or invent results to secure academic papers and the rewards that come with them. This is rightly regarded as misconduct. But there is a new class of bad behaviour — one that is driven by a related but different pressure: ‘impact or perish’.

It is no longer enough for scientists to publish their work. The work must be seen to have an influential shelf life. This drive for impact places the academic paper at the centre of a web of metrics — typically, where it is published and how many times it is cited — and a good score on these metrics becomes a goal that scientists and publishers are willing to cheat for.

Collectively, these new practices don’t seek to produce articles that are based on fraudulent evidence or claims. Rather, they use fraudulent means to secure their publication, enhance their impact and inflate the importance of those who write them. They are on the march — and scientists no longer have to look far to find them. News about research now includes regular reports of authors who supply fake e-mail addresses of suggested peer reviewers. They then use those addresses to offer reports that are supportive enough to ensure that the paper is published. ‘Review and citation’ rings go a step further, trading favourable fake reviews for citations to the reviewer’s work. Others hack publisher databases to seek more invitations to review papers, and so possibly insert more citations to their own work.

All metrics of scientific evaluation are bound to be abused. Goodhart’s law (named after the British economist who may have been the first to announce it) states that when a feature of the economy is picked as an indicator of the economy, then it inexorably ceases to function as that indicator because people start to game it.

What we see today, however, is not just the gaming of science metrics indicators, but the emergence of a new kind of metrics-enabled fraud, which we can call post-production misconduct. It seems to be as widespread as other forms, with at least 300 papers already retracted because the peer review had been tampered with.

A curious feature of this kind of misconduct is that the work itself — the science reported in the paper — is usually not in question. Those responsible for this kind of post-production misconduct seek to extract value not from the article itself, but from its citations. From their point of view, it does not matter whether the article is ever read by a scientist, only that its citations will be harvested by bots.

This means that unlike data fraud and other forms of conventional misconduct, post-production misconduct does not necessarily pollute the scientific record with false results. But it does erode the credibility of the publication system. And it is more common in

emerging countries, perhaps because universities there place the most emphasis on metrics to quickly become globally visible.

How can it be tackled? Post-production misconduct is less likely to be the work of individuals — a hyper-productive protégé operating under the protection of an established mentor who is unwilling to ask too many questions — but increasingly emerges from collaborations. As such, its traces are usually beyond the reach of peer review, which itself is often targeted by these fraudulent schemes.

The exposure of citation and peer-review rings has generally been down to data analysis — of the wording of reviews, review turnaround times, citation patterns and the mutual relationships between authors and reviewers across different publications. Much of this can be mined only by teams of investigators who are carefully poring over journal databases. But publishers consider this type of information proprietary, so when irregularities are found and journals retract articles, they typically offer little detail. After all, these investigations expose weaknesses in their systems and services. (That’s why the new breed of grass-roots watchdogs such as Retraction Watch and PubPeer are so important.)

Given the increasing awareness of post-production misconduct — and how it undermines the assessment of publicly funded research — funders, policymakers and the science community should ask publishers to make available more of the information needed to investigate it.

The community must realize that, unlike previous fraudsters, from the unknown hoaxer who

planted a mixture of bones in a British gravel pit at Piltown to that of Paul Kammerer, who is blamed for inking features onto the feet of midwife toads to support Lamarckianism evolution, academic misconduct is no longer just about seeking attention. Many academic fraudsters aren’t aiming for a string of high-profile publications. That’s too risky. They want to produce — by plagiarism and rigging the peer-review system — publications that are near invisible, but can give them the kind of curriculum vitae that matches the performance metrics used by their academic institutions. They aim high, but not too high.

And so do their institutions — typically not the world’s leading universities, but those that are trying to break into the top rank. These are the institutions that use academic metrics most enthusiastically, and so end up encouraging post-production misconduct. The audit culture of universities — their love affair with metrics, impact factors, citation statistics and rankings — does not just incentivize this new form of bad behaviour. It enables it. ■

Mario Biagioli is a professor of law and of science and technology studies at the University of California, Davis.
e-mail: mbiagioli@ucdavis.edu

ALL
METRICS
OF SCIENTIFIC
EVALUATION
ARE BOUND TO BE
ABUSED.

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

NANOTECHNOLOGY

Synthetic vaccines fight infection

Nanoparticle-encased vaccines can fend off lethal pathogens in animals, and could allow for a swift response to disease outbreaks.

Vaccines made from live virus can elicit long-lasting immunity, but most are slow and laborious to make. Daniel Anderson at the Massachusetts Institute of Technology in Cambridge and his colleagues instead made a fully synthetic vaccine by encasing antigen RNA in a modified-polymer nanoparticle that protected the RNA from degradation.

Single injections of such vaccines against the Ebola virus and the influenza virus, as well as the parasite *Toxoplasma gondii*, were sufficient to generate immune responses in mice and to protect the animals against otherwise lethal doses of the pathogens.

Proc. Natl Acad. Sci. USA
<http://doi.org/bk45> (2016).

EVOLUTION

Double disaster killed dinosaurs

The extinction of the dinosaurs some 66 million years ago was triggered by massive volcanic eruptions in India, and cemented by the arrival of the Chicxulub meteorite in Mexico.

The two events occurred within a few hundred thousand years of one another, making it difficult to tell which drove three-quarters of life on Earth — including the dinosaurs — to extinction. Sierra Petersen at the University of Michigan in Ann Arbor and her co-workers studied a fossil record of mollusc shells from Seymour Island in Antarctica, and used isotope analysis to estimate past temperature

change. An 8°C spike marking the onset of the Deccan Traps volcanic eruptions is followed by a smaller spike some 150,000 years later, coinciding with the Chicxulub impact.

The scientists conclude that the meteorite delivered the final blow to ecosystems already weakened by the eruptions.

Nature Commun. 7, 12079 (2016)

OCEAN SCIENCE

Ice extent changes with the wind

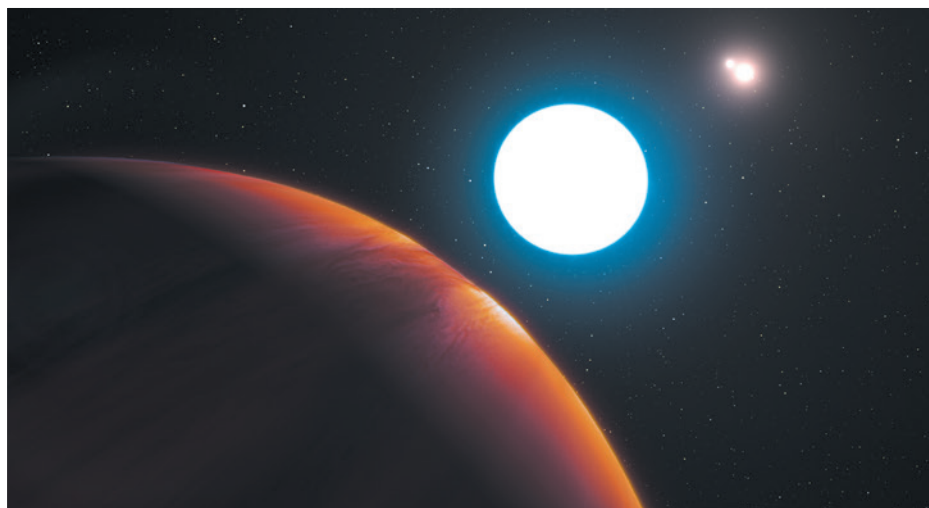
The rate and pattern of summer sea-ice retreat in the Arctic vary markedly from year to year and are driven by multiple atmospheric trends.

Amanda Lynch of Brown University in Providence, Rhode Island, and her team compared atmospheric circulation and Arctic sea-ice trends from 1979 to 2014. Mild summers with warm winds out of Canada and Alaska were associated with marked sea-ice retreat. During cooler summers, warm southerly winds from Siberia tended to favour sea-ice accumulation in the Beaufort Sea and the Canadian Archipelago. In other years, cold winds from the north continually replenished melting sea ice along the southern fringes of the Arctic Ocean.

Understanding how sea ice responds to atmospheric

Few exoplanets have been imaged directly, but the scientists took a series of pictures of a gas-giant planet roughly four times the mass of Jupiter, orbiting the brightest of the three stars. Simulations suggest that it is in a wide, stable orbit, rather than being in the process of being tossed out of the system by gravitational interactions between the stars.

Science <http://doi.org/bk47> (2016)



PLANETARY SCIENCE

Triple star hosts stable planet

An extrasolar planet in an exotic triple-star system lives in surprising harmony with the three stars hanging in its sky.

Kevin Wagner of the University of Arizona in Tucson and his colleagues used the European Southern Observatory's Very Large Telescope in Chile to study the star system HD 131399, which lies about 98 parsecs from Earth in the constellation Centaurus.

patterns might help navigation and exploration in the warming Arctic.

J. Geophys. Res. Atmos.
<http://doi.org/bksf> (2016)

HEALTH

Poor child growth cements poverty

Stunted childhood growth in developing countries results in educational deficits and substantial economic losses.

Using published data on early-childhood growth and the financial benefits of education for 137 developing nations, Günther Fink at the Harvard T. H. Chan School of Public Health in

ESO/L. CALÇADA/M. KORNMESSER

Boston, Massachusetts, and his colleagues calculated the effects of growth delays on children born in 2010. About one-third of the 122.9 million children experienced stunted growth, resulting in the overall projected loss of half a year of educational attainment per child. Globally, early growth delays were projected to cost a total of US\$176.8 billion in lost income each year.

If growth delays were eliminated worldwide, the authors estimated, India would benefit the most, with a projected economic gain of \$37.9 billion per year.

Am. J. Clin. Nutr. 104, 104–112 (2016)

ROBOTICS

Robotic stingray follows the light

A miniature robotic stingray powered by rat heart-muscle cells can swim in a physiological salt solution, guided by light.

Kit Parker at Harvard University in Cambridge, Massachusetts, and his collaborators created the 16-millimetre-long mock-up, weighing just 10 grams, by encasing a gold skeleton mimicking a stingray's shape in an elastic polymer (**pictured**). The stingray can be remotely controlled thanks to a light-activated muscle layer, made from genetically engineered rat cells, which responds to light-pulse frequency.

The artificial fish represents a step towards the development

of adaptive robots with capabilities inspired by nature, the researchers say.

Science 353, 158–162 (2016)

CHEMISTRY

Cosmic rays breed organics in space

Cosmic rays help to form the Universe's complex organic molecules — the building blocks of life on Earth.

The interstellar gas clouds that give birth to stars and planets are rich in organic molecules, but scientists have struggled to explain how these formed. A team led by Ralf Kaiser at the University of Hawaii at Manoa in Honolulu looked into it in the lab, using supercold ices to approximate conditions found in space. They showed that cosmic rays can trigger previously unknown chemical reactions that form sugars and other organic compounds in ice as cold as 10 kelvin. These compounds could sublime into the gas clouds that become stellar nurseries, and eventually lead to the formation of biological molecules.

The discovery should help to explain the origin and evolution of the molecular make-up of the Universe. *Proc. Natl Acad. Sci. USA* <http://doi.org/bk46> (2016)

PRIMATOLOGY

Ancient monkeys used stone tools

Worn rocks may have been used by Brazilian bearded capuchin monkeys hundreds of years ago, marking the earliest evidence for stone-tool use by a species other than humans or chimpanzees.

The oldest stone tools made by ancient human relatives date to 3 million years ago, but the archaeological record for tool use among other primates is scant. Bearded capuchins (*Sapajus libidinosus*; **pictured**) are known to wield rock anvils to rid cashew nuts of their foul-tasting skin. To determine the antiquity of this behaviour,



Michael Haslam at the University of Oxford, UK, and his colleagues excavated a small area in a Brazilian national park populated by capuchins known to use stone anvils.

They found 69 buried stones with surface cut marks similar to those seen on anvil stones. Carbon dating suggested that some of the tools were buried as many as 750 years ago. *Curr. Biol.* 26, R515–R522 (2016)

DEVELOPMENTAL BIOLOGY

Mum's diet affects offspring's genes

Poor nutrition during pregnancy stunts the growth of young mice by modifying their gene expression.

Michelle Holland and Vardhman Rakyen at Queen Mary University of London and their colleagues fed female mice diets containing either 8% or 20% protein throughout pregnancy and until weaning. They analysed patterns of methylation — which can influence gene expression — on the DNA of the rodents' offspring.

Pups from mothers fed the low-protein diet were, on average, 25% smaller at weaning. This effect was further influenced by variation within an animal's many gene copies for ribosomes, the cell's protein-construction machines. The extent of growth restriction depended on the proportion an individual had of a particular gene variant.

Studying the effects of methylation and other chemical marks on ribosomal genes may shed light on some human diseases, the authors say. *Science* <http://doi.org/bk5b> (2016)

METABOLISM

Sweetness–energy mismatch

Regular consumption of the artificial, calorie-free sweetener sucralose causes animals to overeat.

Greg Neely at the University of Sydney in Australia, Herbert Herzog at the Garvan Institute of Medical Research, also in Sydney, and their colleagues fed fruit flies a sucralose-enriched diet for between one and six days. After five days or longer on the sucralose diet, the flies ingested up to 30% more calories than did flies fed a normal diet throughout. The insects also became hyperactive and glucose intolerant.

The group discovered that a neural circuit was activated in the brains of sucralose-fed flies that is normally switched on during fasting, and is known to control hunger and sweet-taste intensity. They found similar responses in mice fed sucralose. *Cell Metab.* <http://dx.doi.org/10.1016/j.cmet.2016.06.010> (2016)

NATURE.COM

For the latest research published by Nature visit:

www.nature.com/latestresearch



SEVEN DAYS

The news in brief

RESEARCH

Dwarf planet find

An international hunt for objects in the distant Solar System has turned up a new dwarf planet far beyond Neptune. The object, named 2015 RR245, is about 700 kilometres wide and is on the way to making its closest approach to the Sun, at 5 billion kilometres, or 34 times the Earth–Sun distance. At its farthest point, the object will be some 19 billion kilometres from the Sun, giving it one of the largest orbits of any known dwarf planet. Astronomers with the Outer Solar System Origins Survey discovered 2015 RR245 during their ongoing search at the Canada-France-Hawaii Telescope in Hawaii.

Trial deaths

A clinical trial run by cancer-drug developer Juno Therapeutics of Seattle, Washington, has been put on hold by regulators following the deaths of two patients in the same week, the company announced on 7 July. The trial was studying the effects of genetically engineered T cells, a type of immune cell, on a form of leukaemia. Both patients died after a chemotherapy drug was added to the treatment. The two latest deaths follow a similar one in May.

Wellcome opens up

The Wellcome Trust, one of the world's largest biomedical charities, will launch its own open-access publishing venture later this year. The venture, called *Wellcome Open Research*, was announced by the trust on 6 July. It will be managed by *F1000Research*, a platform that publishes manuscripts within days of submission, after a quick sanity check by its in-house

editors; peer review is then done after publication. Unlike many journals, authors choose their own reviewers, and reviews are generally short, averaging 400 words. Wellcome thinks that other major research funders are likely to create similar systems, and that over time the ventures could merge into one big, international open-access publishing platform. See go.nature.com/29nth06 for more.

Olympians study

Researchers in the United States hope to enrol 1,000 US Olympic athletes and staff in a Zika virus study. The participants are travelling to Rio de Janeiro, Brazil, for the Olympic and Paralympic games in August and September. The

mosquito-borne virus, which has been linked to birth defects, has spread across Brazil and the Americas in recent months. The study, announced on 5 July and funded by the National Institutes of Health, will examine factors such as where the virus resides in the body and whether birth defects result from infection. Participants will be asked to complete health questionnaires and to provide samples of bodily fluids.

Zika vaccine deal

French drugmaker Sanofi announced on 6 July that it has signed an agreement with the Walter Reed Army Institute of Research, a laboratory run by the US Department of Defense, to develop a vaccine for the Zika virus. Walter Reed plans to share its inactivated

virus vaccine technology with Sanofi's vaccine division, which will handle clinical development and regulatory approval. The partnership hopes to produce a vaccine ready for human trials starting in October. Vaccines for the mosquito-borne virus, which is spreading through the Americas and has been linked to birth defects, have proved challenging to develop.

PEOPLE

French agency row

A ferocious battle has erupted in France over who will head the country's National Institute for Agricultural Research (INRA) for the next four years. The scientific community overwhelmingly backs the incumbent, François Houllier,



JAMES COOK UNIVERSITY

Huge mangrove die-off in Australia

Mangrove forests in northern Australia are dying off on a large scale, reveal satellite and aerial images released on 11 July. Norm Duke, a mangrove scientist at James Cook University in Townsville, Australia, says that 7,000 hectares of mangroves have been affected along the Gulf of Carpentaria. The die-off seems to be

connected to extreme ocean warming associated with the 2015–16 El Niño event in the tropical Pacific Ocean, which also led to coral bleaching in the Great Barrier Reef. Calling the event “unprecedented”, Duke says that scientists must investigate the cause and scale up monitoring efforts.

for a second term, and rejects President François Hollande's nominee Philippe Mauguin, chief of staff to the agriculture minister. The vast majority of INRA's researchers have signed an online petition that claims that Mauguin is a political choice, citing his insufficient scientific experience; most opposition parliamentarians have called for Hollande to suspend the nomination process. As *Nature* went to press, the National Assembly and Senate economic affairs committees were still scheduled to vote on Mauguin's appointment on 13 July.

Scientist acquitted

A judge in Verona, Italy, has acquitted virologist Ilaria Capua of all charges of attempting to create and profit from an epidemic of avian influenza. In the 5 July ruling, the judge also dismissed related lesser charges because the time limit had passed for legal proceedings to be brought. The accusations had been under investigation for more than a decade. Capua (pictured), former research director of the Experimental Institute for the Prevention of Animal Diseases in Legnaro and one of 16 defendants in the complex case, took a job as director of the Center for Excellence in One Health Research at the University of



Florida in Gainesville in June, citing the difficulties of doing science in Italy.

BUSINESS

Theranos troubles

Elizabeth Holmes, founder and chief executive of the beleaguered blood-analysis firm Theranos in Palo Alto, California, has been banned by US regulators from running a lab for two years and fined an unspecified amount. The 7 July announcement is the latest misfortune for the Silicon Valley company, which rose to prominence with promises of a technology that could perform a wide variety of diagnostic tests from a few drops of blood. But its bold claims faced scepticism and government scrutiny following investigations last year.

Web of Science sale

Major news and information company Thomson Reuters announced on 11 July that

it will sell its intellectual-property and science division, including its subscription-based Web of Science citation databases, to two private-equity firms for US\$3.55 billion. The company has long been expected to part with the division, which employs some 3,200 people, to focus on its core businesses of financial, trade and regulatory information. Industry observers expect the equity firms — Onex Corporation and Baring Private Equity Asia — to break up the division and resell the parts at a profit.

POLICY

Care.data killed off

The UK health department announced on 6 July that it is abandoning care.data, a controversial programme of the National Health Service in England that would have centralized medical records and potentially made them available to researchers. Public misgivings over the protection of patient information delayed the launch of care.data in 2014, and in 2015 the UK government launched two independent reviews of the programme. Those reviews, also released on 6 July, called for new data-security standards and proposed a different model for how

COMING UP

11–20 JULY

The United Nations holds a High-Level Political Forum on Sustainable Development in New York City.
go.nature.com/29h7y7t

16 JULY

A Russian Soyuz rocket launches from Baikonur Cosmodrome in Kazakhstan to deliver cargo to the International Space Station.

patients can consent to, or opt out of, sharing their health records.

EVENTS

Brexit woes abound

UK researchers continue to struggle with the uncertainty created by the country's 23 June vote to leave the European Union. Theresa May, former Home Secretary who was confirmed this week as Britain's new prime minister, said that the future right to residence for EU nationals working in the United Kingdom is not guaranteed. This includes 15% of the academic workforce. In other news, the Russell Group of leading UK universities expressed concern on 5 July about racism since the referendum. And life-sciences minister George Freeman has enlisted two drug-company bosses, GlaxoSmithKline's Andrew Witty and AstraZeneca's Pascal Soriot, to co-chair a Life Science Steering Group with him. It will produce recommendations on how to deal with post-referendum issues. See go.nature.com/29slyz for more.

NATURE.COM

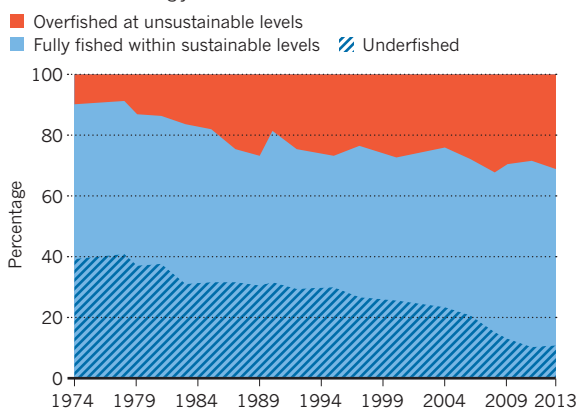
For daily news updates see:
www.nature.com/news

TREND WATCH

The world is eating more fish than ever — with farmed, rather than wild-caught, animals driving the increase, reveals a report from the Food and Agriculture Organization of the United Nations, out on 7 July. In 2014, people ate an average of 20.1 kilograms of fish, up from 9.9 kg per year in the 1960s. But wild fish stocks are also being increasingly overexploited. In 1974, 90% of wild stocks were being fished sustainably; by 2013, this had declined to 68.6%. See go.nature.com/29cvzf for more.

THE SINKING SUSTAINABILITY OF FISH STOCKS

Despite good practice in some areas, the world's marine fish stocks are increasingly overfished.



NEWS IN FOCUS

TECHNOLOGY Mobile-phone expansion threatens weather satellites **p.208**

NEUROSCIENCE Data gold mine could reveal how neurons compute **p.209**

PUBLISHING Tide turns against controversial impact factor **p.210**



EVOLUTION Chinese fossils challenge ideas about human origins **p.218**

MUJAHID SAFODIEN/AFP/GETTY



Student protests at universities have thrust South Africa's research into an unwelcome spotlight.

RESEARCH AND SOCIETY

South African universities awash in political turmoil

Campus violence around the country has taken a toll on research activities.

BY ERIKA CHECK HAYDEN

South Africa's universities have long been viewed as among the strongest in sub-Saharan Africa. But ongoing financial and political turmoil are endangering research at the nation's centres of higher learning.

It is an extension of the overall crisis that currently faces South Africa. The country's economy is in tatters, with its currency falling in value, and a corruption scandal has weakened President Jacob Zuma and his ruling political party. The upheaval is an unwelcome backdrop to the upcoming 21st International

AIDS conference in Durban on 18–22 July (see page 214).

Funding cuts have threatened agencies such as the South African Medical Research Council (MRC), and have harmed researchers' abilities to run their labs and recruit young scientists. They have also delayed planned projects at South African universities. The MRC is fighting a 7% budget cut for 2017–18, says council president Glenda Gray, and probably more cuts the year after. "This would be devastating," she says. "By cutting off science funding, you cut off your ability to be globally competitive."

But the true effects have not yet been felt,

says vaccine researcher Shabir Madhi, executive director of South Africa's National Institute for Communicable Diseases in Johannesburg. He worries that young scientists' careers will sputter as a result of funding shortfalls, because junior researchers depend on government and university money to launch their programmes. Long-established scientists draw much of their research funding from outside South Africa and are less affected by the cuts.

Attempts by universities to compensate for slashes to funding have been challenged, sometimes violently. A proposed 10.5% fee hike at the University of the Witwatersrand ►

► in Johannesburg last October sparked a student protest movement, known on social media as the “#FeesMustFall” campaign. The protests spread to other universities, which had also proposed fee increases of 10–12%. The movement successfully squashed the hikes, and protestors are now pressing for free tuition at all South African universities.

Without fee increases, however, universities are facing immediate budget shortfalls, which have forced them to make widespread cuts. Astronomer David Block at the University of the Witwatersrand says that he and his colleagues were told at a faculty meeting last month to save money by cutting their use of water, heat and electricity. Earlier this year, he attempted to recruit a promising postdoctoral researcher, but failed because his department lacked money for new hires. “It really has reached a crisis — we’re under tremendous strain.”

UP IN SMOKE

Projects such as a programme to train vaccinology researchers at various institutions, including the University of the Witwatersrand, have had to find outside funding — a temporary stopgap.

Researchers are also worried about access to infrastructure, ever since protests ahead of municipal elections on 3 August led to campus vandalism. Unrest in January shut down University of Pretoria campuses for weeks, and in February, the Potchefstroom campus of North-West University was closed after students torched administration buildings, including a science centre. In May, arsonists burned down a historic auditorium at the University of Johannesburg.

Researchers and university administrators worry that political violence is becoming a new normal. During some of the worst mayhem, in May, Alta Schutte, director of the hypertension unit of North-West University said, “When I go home every day, I am a bit concerned that when I come back, my office, my lab, my hypertension clinic or my biobank will not be there.”

The protest movement is a response to the nation’s persistent inequality. “An upper-middle-income country like South Africa should widen access to education,” says cardiologist Bongani Mayosi at the University of Cape Town. But the violence and intimidation have gone too far, he says.

Danie Visser, deputy vice-chancellor for research and internationalization at the University of Cape Town, agrees. “We are probably at a critical juncture: if the country is able meaningfully to address the issues that brought about the student protests in the first place, our universities — and therefore also our research — will survive and flourish.” ■



Severe-weather forecasts deteriorate when wireless broadband interferes with satellite transmissions.

METEOROLOGY

Interference puts satellite data at risk

US plan to expand mobile-phone bandwidth raises alarm.

BY ALEXANDRA WITZE

As Hurricane Patricia barrelled down on Mexico last October, forecasters at the US National Oceanic and Atmospheric Administration (NOAA) grabbed as many satellite images as they could to track its progress. But at least one crucial shot failed to download. A 22 October image from the Geostationary Operational Environmental Satellite (GOES) system showed a black swathe — no data — across most of the Pacific Ocean.

“You couldn’t even see the hurricane,” says Al Wissman, chief of data management and continuity operations for NOAA’s satellite and information service in Silver Spring, Maryland. “That’s how devastated the imagery was.”

The culprit was radio interference from mobile-phone companies. And the problem may soon get worse. The US Federal Communications Commission (FCC) is considering whether to allow a commercial mobile-phone company to share a crucial, additional set of frequencies that NOAA uses for time-critical weather transmissions.

If the application is granted, Ligado Networks of Reston, Virginia, will begin transmitting at frequencies between 1,675 and 1,680 megahertz. That overlaps with the communications range of NOAA’s next

generation of GOES satellites, starting with the game-changing GOES-R probe that is set to launch in November. The satellite will transmit in three times the number of channels as do current satellites, providing images with four times the current resolution, and it will scan for weather events five times faster.

Last month, emergency managers, pilots, private weather forecasters and other groups flooded the FCC with letters arguing against the change. Researchers will discuss the proposal at a meeting of the American Meteorological Society (AMS) in Tuscaloosa, Alabama, on 21 July.

Wireless broadband has been a boon for meteorologists, who can now send crucial tornado, hurricane and other alerts directly to people’s smartphones. “But it can’t come at the risk of interrupting important weather communications that are used in order to be able to deliver the most accurate and reliable forecast,” says Jonathan Porter, vice-president of innovation and development at the private forecasting company AccuWeather in State College, Pennsylvania. Porter also chairs an AMS committee on spectrum allocation.

In general, the US government sets aside swathes of radio frequencies for purposes that protect safety and national security, such as weather forecasting. But in 2010, President Barack Obama told the various agencies

RICKY CARLTON/WASHINGTON POST

that regulate spectrum-sharing to free up 500 MHz for wireless broadband use by 2020. In November 2012, a company that later evolved into Ligado filed a request to share the 1,675–1,680-MHz band.

Commercial mobile-phone companies are already transmitting at slightly lower frequencies, the 1,670–1,675-MHz band — a situation that has caused problems with NOAA data.

In a representative sample of GOES imagery taken between May and September 2015, the agency found that 3.6% of the data during that stream had been subject to interference. And in May of this year, NOAA clocked 30 events in which satellite transmissions had dropped out, either streaking or nearly obliterating the images. “We consider that to be unacceptable,” Wissman says.

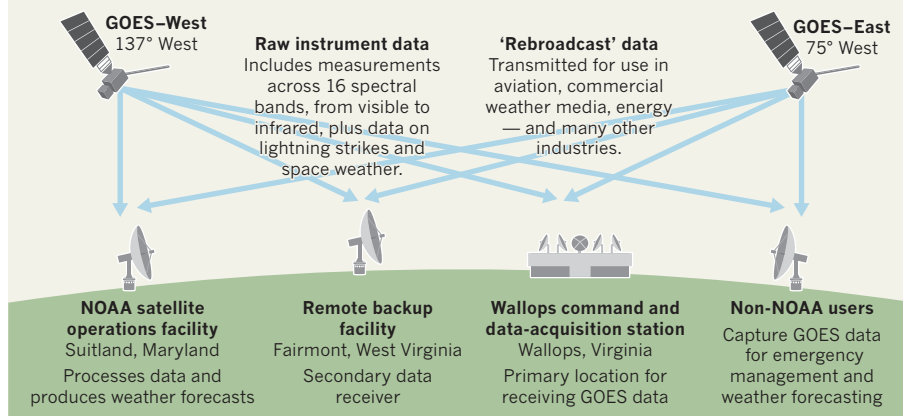
In response, NOAA has begun to shift the transmission frequencies for the radiosonde balloons it launches to obtain vertical profiles of the atmosphere. It also redesigned an aspect of its GOES-R transmissions to be centred on 1,686.6 MHz, in the hope that this would be high enough to escape the interference.

But that change affects only how GOES-R relays its own imagery to Earth. GOES-R has a second job as a sort of internet in the sky, relaying data from 27,000 ground stations including stream gauges, tsunami buoys and seismic stations (see ‘Weather watchers’). If Ligado’s application is granted, that ‘rebroadcast’ service is likely to be interrupted — affecting forecasts of phenomena such as the spread of smoke during wildfires or the disruption of plane flights by volcanic ash.

“It’s just an untenable situation to have in a critical situation,” says William Mahoney, an atmospheric scientist at the National Center for Atmospheric Research in Boulder,

WEATHER WATCHERS

The US government’s Geostationary Operational Environmental Satellite (GOES) system monitors atmospheric and surface conditions in the continental United States — collecting data that power the country’s weather forecasts.



Colorado, and head of the AMS commission on the weather, water and climate enterprise.

One of GOES-R’s big advantages is that it will send updated data as often as every 30 seconds. That’s much more frequent than the 10–30-minute refresh time of the current GOES series, so any disruption to the real-time data flow will be much worse, Porter says.

Ligado has proposed ways to address the concerns, such as establishing blackout zones around NOAA’s receiving stations or creating a cloud-based computing network to handle data distribution for non-NOAA users.

But many of those who have commented publicly are sceptical about such plans. The World Meteorological Organization pointed out that cloud computing is vulnerable when weather data are most needed: during severe storms.

The FCC is accepting replies to the original set of public comments until 21 July. After that, it will grind slowly towards a decision.

In Tuscaloosa next week, meteorologists will sit down for a public discussion with representatives from Ligado about the best way forward. Porter, who will chair the panel, hopes that the government will proceed slowly — perhaps by delaying the bandwidth-sharing or at least phasing it in slowly and documenting any interference.

“This is not just, ‘Oh, a few weather forecasts,’” says Renee Leduc Clarke, a consultant with Narayan Strategy in Washington DC who has been working with clients on the spectrum-sharing issue. “This is equal to lives and property inside our economy — the same economy we’re trying to boost with wireless broadband.” ■

NEUROSCIENCE

Brain-data gold mine released

Massive survey of mouse visual-cortex activity aims to reveal brain’s computational rules.

BY HELEN SHEN

Inspired by the large-scale sky surveys with which astronomers explore the cosmos, neuroscientists in Seattle, Washington, have spent four years systematically surveying the neural activity of the mouse visual cortex. The Allen Brain Observatory’s first data release, on 13 July, provides a publicly accessible data set of unprecedented size and scope, designed to help scientists to model and understand the human brain.

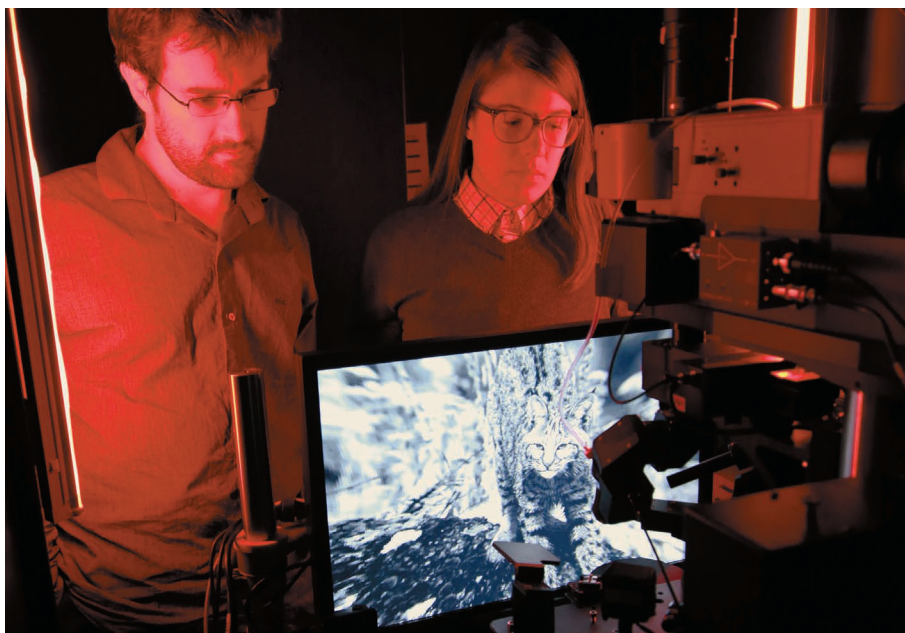
The project is part of an ambitious ten-year brain-research plan announced in 2012 by the

Allen Institute for Brain Science. Designed to catalogue neurons and their electrical characteristics in minute detail, the initiative aims to enable new insights into how perception and cognition arise.

To compile the brain observatory’s first data set, researchers used a specialized microscope to record calcium waves that occur when neurons fire, sampling activity in 25 mice over 360 experimental sessions, while the animals viewed a battery of visual stimuli such as moving patterns of lines, images of natural scenes and short movies. The data set so far includes 18,000 cells in 4 areas of the visual cortex,

making it one of the largest and most comprehensive of its kind. The set also includes information about each neuron’s location and its expression of certain genetic markers. At 30 terabytes, the raw data are too large to share easily, but users can download a more manageable processed data set, or explore it online.

“It’s amazing,” says Anne Churchland, a neuroscientist at Cold Spring Harbor Laboratory in New York. “There’s no other effort I know of where people have looked at so many brain areas with so many stimuli — and importantly, where the data are freely available as well.” ▶



Researchers monitored mice's brain activity while the animals viewed images such as this cat.

► Other labs have collected similar data, but on a much smaller scale, with fewer animals or fewer neurons. This information has been difficult to merge and compare, as a result of differences in the species, techniques or brain regions examined. And most data remain in the hands of individual labs.

To create the unusually extensive Allen data set, more than 100 researchers developed and used standardized equipment and protocols for every stage of the experiment. This allowed them to repeatedly and systematically sample the same populations of neurons across many animals and sessions.

Now, Allen Institute researchers plan to

monitor activity while the mice carry out behavioural tasks. The scientists also want to use more recording techniques, and to extend their sampling across the entire mouse visual cortex and beyond. Christof Koch, president of the Allen Institute, hopes that over the next 3–4 years, the project will evolve into a true observatory, with researchers able to request certain experiments — the results of which will be made publicly available.

The project's neural-activity map could help to fill out a picture of what cell types live in the brain and how they work together. Ultimately, the Allen Institute wants its own researchers and others to be able to use the massive data set

to help to uncover the fundamental computational principles that underlie cognition. This lofty goal is shared by the US government's Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative, which was launched in 2013 with the Allen Institute among its private partners. But whereas the BRAIN Initiative has largely supported individuals and small groups of investigators with conventional grants, the Allen Institute has concentrated personnel and money on a small number of large projects. It aims to create public research tools that would be unfeasible for individual labs to produce.

Armed with a sweeping survey of neural activity, Koch says, theoreticians will be able to design more accurate models of brain function, and find better ways to test the validity of existing models. But he is also realistic about the challenges ahead. "We're under no illusions that now we have all this data that the solution will jump out at us," says Koch.

The Allen Brain Observatory's impact will depend in part on whether the neuroscience field embraces this experiment in communal research. Early reactions suggest that researchers are eager to participate. Churchland says that the in-depth information about how different visual areas respond to stimuli could help to guide and fine-tune her experiments. The data could also help labs that lack access to highly specialized imaging equipment, she adds.

Theoreticians, too, are looking forward to delving into the data. "This is basically a bonanza," says computational neuroscientist Steven Zucker at Yale University in New Haven, Connecticut. "It's as if somebody opened the door into the world's biggest neuroscience lab for theoreticians around the world and said, 'Come on in and play with our data.'" ■

ALLEN INST.

BIBLIOMETRICS

Publishing elite turns against impact factor

Senior staff at societies and leading journals want to end inappropriate use of the measure.

BY EWEN CALLAWAY

The tide is turning against the impact factor — one of the publishing industry's most contentious metrics — and its outsized impact on science.

Calculated by various companies and promoted by publishers, journal impact factors (JIFs) are a measure of the average number of citations that articles published by a journal

in the previous two years have received in the current year.

They were designed to indicate the quality of journals, but researchers often use the metrics to assess the quality of individual papers — and even, in some cases, their authors.

Now, a paper posted on the preprint server bioRxiv on 5 July, authored by senior employees at several leading science publishers (including *Nature's* owner, Springer Nature),

calls on journals to downplay the figure in favour of a metric that captures the range of citations that a journal's articles attract (V. Larivière *et al.* Preprint at bioRxiv <http://doi.org/bmc2>; 2016).

And in an editorial that will appear on 11 July in eight of its journals, the American Society for Microbiology (ASM) in Washington DC will announce plans to remove the impact factor from its journals and website, as well as from

marketing and advertising.

"To me, what's essential is to purge the conversation of the impact factor," says ASM chief executive Stefano Bertuzzi, a prominent critic of the metric. "We want to make it so tacky that people will be embarrassed just to mention it."

Bertuzzi was formerly the executive director of the American Society for Cell Biology, which banned the mention of impact factors from its annual meeting.

BRACE FOR IMPACT

Heidi Siegel, a spokesperson for London-based business-analytics firm Thomson Reuters, the major publisher of the JIF, says that the measure is a broad-brush indicator of a journal's output — and should not be used as a proxy for the quality of any single paper or its authors. "We believe it is important to have a measure of the impact of the journal as a whole, and this is what the JIF does," says Siegel.

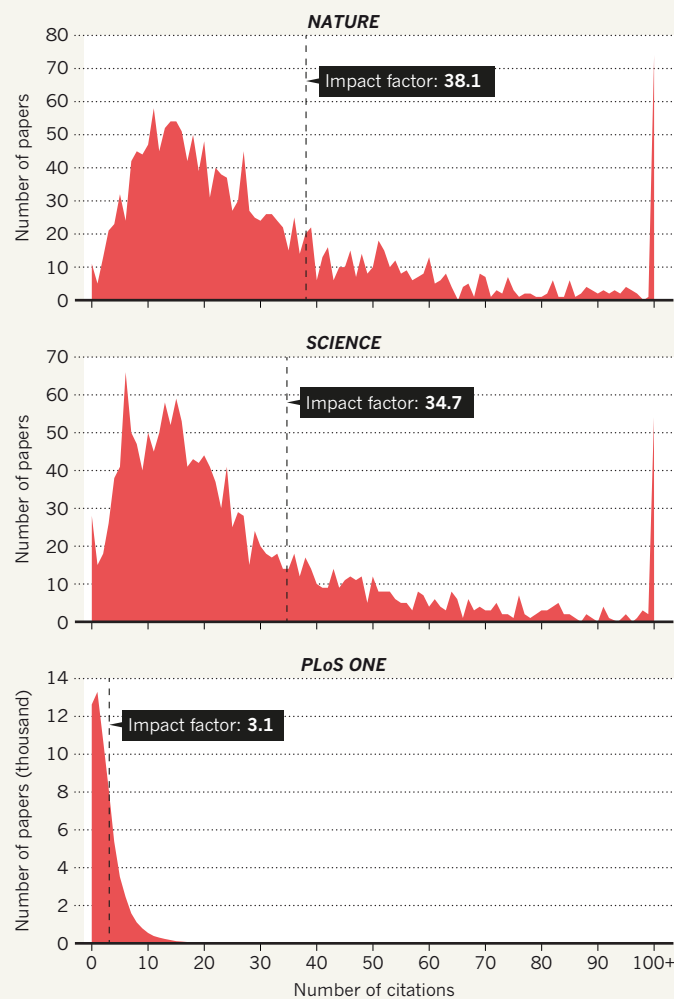
But many scientists, funders and journals do not use it that way, notes Stephen Curry, a structural biologist at Imperial College London who is lead author of the latest paper. Many researchers evaluate papers by the impact factor of the journals in which they appear, he says, and the metric can also influence decisions made by university hiring committees and funding agencies.

Past research suggests that such uses are inappropriate. To emphasize some limitations of the JIF, Curry's team plotted the distribution of citations for articles published in 2013–14 in 11 journals, including *Science*, *Nature*, *eLife* and 3 Public Library of Science (PLOS) journals (see 'The impact factor's long tail'). These citations are used to calculate the 2015 impact factors.

Curry's co-authors include senior employees at Springer Nature, *eLife*, PLoS, the Royal Society (which publishes several journals) and EMBO Press, as well as Marcia McNutt, who

THE IMPACT FACTOR'S LONG TAIL

Journal impact factors are influenced heavily by a small number of highly cited papers. For all journals analysed, most papers published in 2013–14 garnered many fewer citations than indicated by the impact factor.



stepped down on 1 July from her role as editor-in-chief of *Science*.

Most of the papers garnered fewer citations than the impact factor for their journal: 74.8% of *Nature* articles were cited below its impact factor of 38.1, and 75.5% of *Science* papers were cited fewer than 35 times in 2 years (its impact factor was 34.7). *PLoS Genetics* had the lowest proportion of papers with fewer citations than its impact factor of 6.7, at 65.3%.

Highly cited papers explain this disconnect. *Nature*'s most cited paper in the analysis was

papers to read, for instance.

"Denying the value of impact factors in this situation essentially means that we deny the value of the entire journal publishing system and of all the work done by journal editors and peer reviewers to carry out quality control," Waltman says. "To me, this doesn't make sense."

Anti-impact-factor crusaders say that it will take time to diminish the influence of the figure, let alone exile it. "This is a cultural thing," says Bertuzzi, "and it takes pressure from multiple points to change behaviour". ■

referenced 905 times and *Science*'s 694 times. *PLoS ONE*'s biggest paper accrued 114 citations, compared with its impact factor of 3.1.

Some journals, such as those published by the Royal Society and EMBO Press, already publicize citation distribution. Curry and his fellow authors explicitly recommend that other publishers play down their impact factors, and, instead, emphasize citation-distribution curves, such as those that the team generated, because they provide a more informative snapshot of a journal's standing. The preprint includes step-by-step instructions for journals to calculate their own distributions.

A MEASURE OF CHANGE

A spokesperson for *Nature* says that the journal will soon update its websites "to cover a broader range of metrics", and a representative of *Science* has stated that the journal will consider the proposal once the preprint article is published in a peer-reviewed journal.

Ludo Waltman, a bibliometrics researcher at Leiden University in the Netherlands, says that citation distributions are more relevant than impact factors for high-stakes decisions, such as hiring and promotion. But he is wary of doing away with impact factors entirely; they can be useful for researchers who are trying to decide which among a pile of



BREXIT AND SCIENCE



The latest as scientists grapple with the effects of Brexit
go.nature.com/brexit

MORE NEWS

- Strange planet has triple sunsets and a super-long year go.nature.com/29axu5o
- Clouds get high on climate change go.nature.com/29aedsu
- Dead X-ray satellite reveals galaxy cluster surprise go.nature.com/29foamf

NATURE PODCAST



The psychological toll of war; how to count the dead; and predicting conflict
nature.com/nature/podcast

RESEARCH COMMUNITY

Self-citation rates higher for men

The trend has taken off over the past 20 years.

BY DALMEET SINGH CHAWLA

Men are more likely than women to cite their own papers, according to an analysis of 1.5 million studies published between 1779 and 2011 — although the reason remains unclear.

The study, led by sociologist Molly King at Stanford University in California, looked at papers across disciplines in the digital library JSTOR. On average, men cited themselves 56% more than women. But over the past two decades, men's self-citation rates rose to 70% more than women, despite an increase of women in academia. Between 1779 and 2011, about 10% of a given paper's references were likely to be self-citations, regardless of the author's gender.

There are some limitations to the conclusions of the study, which was released on 5 July (M. King *et al.* Preprint at <http://arxiv.org/abs/1607.00376>; 2016). Men tend to hold more-senior positions in academia, publish more papers and so have more work to self-cite, says Cassidy Sugimoto, an information scientist at Indiana University Bloomington. So whether this trend is a by-product of the under-representation of women in senior academic positions, or some separate effect, is unclear.

King and her colleagues also deciphered the gender of authors listed on a publication using their first names and their associated sex in US Social Security Administration records. They discarded gender-neutral names and authors listed with only a first initial, which may have disproportionately excluded women, who might be more likely to use their initials to obscure their gender.

The researchers could match a sex to only 56.4% of the authors. "That's a very big portion left out, especially considering that women may be actively hiding their gender," says Adrian Letchford, a data scientist at the University of Warwick in Coventry, UK. Despite these caveats, the trend is still important and must be addressed, says Sugimoto.

Some self-citations could be deliberate attempts to boost an academic's citation count, King says. Hiring and tenure committees should consider this when assessing the impact of a researcher's work, she adds. ■



Renewable power sources such as wind now provide about one-third of Germany's electricity.

RENEWABLE POWER

And now for the energy forecast

Germany works to predict wind and solar power generation.

BY QUIRIN SCHIERMEIER

The rows of towering wind turbines and legions of glistening solar panels spread across Germany's landscape are striking emblems of the country's shift to non-nuclear, low-carbon power. But although Germany is the world's poster child for renewable energy, its grids cannot yet cope with the erratic nature of wind and solar power.

In June, German meteorologists, engineers and utility firms began to test whether big data and machine learning can make these power sources more grid-friendly.

"To operate the grid more efficiently and keep fossil reserves at a minimum, operators need to have a better idea of how much wind and solar power to expect at any given time," says Malte Siefert, a physicist at the Fraunhofer Institute for Wind Energy and Energy System Technology in Kassel, Germany, and a leader on the project, called EWeLiNE.

At about 45,000 megawatts, Germany's

wind-power capacity is the third largest in the world, behind China's and the United States'. And Germany is outperformed only by China in solar capacity. But the pace of the country's switch to renewables and its ambitions are unrivalled. Renewables now provide about one-third of domestic electricity and the government has promised that by 2050, at least 80% of the country's electricity will come from renewables.

The trouble is that on calm and cloudy days, grid operators still need to call on conventional power stations to meet expected demand. And on unusually sunny and windy days — such as on 8 May, when for about 4 hours wind and solar power generated more than 90% of the electricity that Germany consumed — they must swiftly order coal and gas-fired power stations to reduce their output lest an influx of power 'congests' the grid and increases the risk of failures.

Such requests, called re-dispatches, cost German customers more than €500 million

JULIAN STRATENSCHULTE/DPA/PA IMAGES

(US\$553 million) a year because grid operators must compensate utility firms for adjustments to their inputs. They can also lead to needless carbon dioxide emissions if grid operators generate extra power that ends up being wasted. “It is quite a concern that renewable energy here is expanding so fast without a proper data-base for accurate power forecast,” says Renate Hagedorn, a meteorologist with the German weather service in Offenbach.

EYE OF THE STORM

Standard weather models predict the strength and arrival times of storms and weather fronts in a given region. But they cannot, for example, predict wind strength at the hub of a turbine, which determines the amount of power the turbine will produce. The €7-million EWeLiNE project, a collaboration that includes three major grid operators — 50Hertz, Amprion and TenneT — and that is funded by the federal ministry for economic affairs and energy, set out in 2012 to provide load forecasts that are specific to the needs of grid operators.

Most wind turbines are equipped with devices that measure the wind speeds at their hubs, and some solar panels contain sensors for sunlight intensity. EWeLiNE combines these data with other atmospheric observations — from ground-based weather stations, radar and satellites — and sophisticated computer

models predict power generation over the next 48 hours or so. The team checks these power forecasts against what actually materializes, and machine learning then improves the predictive models.

The EWeLiNE researchers began testing their system using solar-panel and wind-turbine data from across Germany last month. Eventually, the idea is for grid operators to use the power

“Operators need to have a better idea of how much wind and solar power to expect at any given time.”

plans to have real-time transmission capabilities for most of the wind and solar facilities in Germany.

There are signs that the approach will work. The National Center for Atmospheric Research (NCAR) in Boulder, Colorado, started on a similar system in 2009, and it is now operational in eight US states. At Xcel Energy, the utility firm with the highest total wind capacity in the United States, the number of forecasting errors has dropped since 2009, saving customers some US\$60 million

forecast to guide these requests. But very few wind and solar facilities are set up to transmit the data in real time, so the results cannot yet be used to adjust how much power is produced. In two years, EWeLiNE

and reducing annual CO₂ emissions from fossil-reserve power generation by more than a quarter of a million tonnes per year, says Drake Bartlett, a renewable-energy analyst with the firm who is based in Denver, Colorado.

“Germany has some very good modellers who are already doing a nice job without the real-time data that we have,” says Sue Haupt, who oversees weather-system research at NCAR. “Once they get access to that data I am sure it’ll be used to great effect.”

EWeLiNE cannot simply use the NCAR system because weather models and the algorithms that convert weather predictions into power forecasts differ between the United States and Germany. ■

CORRECTIONS

The News story ‘Academics across Europe join ‘Brexit’ debate’ (*Nature* **530**, 15; 2016) should have made it clear that Brexit would mean that the United Kingdom would have to renegotiate access to ITER but that it wouldn’t necessarily go to the back of the queue. And the News story ‘Zika raises wider birth-defect issue’ (*Nature* **535**, 17; 2016) wrongly said that Stanley Plotkin has retired, he is still working at the University of Pennsylvania in Philadelphia.



Access to antiretroviral drugs has improved the quality of life of millions of people in South Africa: this woman could not walk before receiving treatment.

A NEW ERA FOR HIV

South Africa has developed the biggest programme of antiretroviral therapy in the world. Now scientists are exploring the long-term consequences of the drugs.

BY LINDA NORDLING

Sixteen years ago, a sickly eleven-year-old became the human face of the AIDS epidemic that was sweeping South Africa. Standing up to speak in front of thousands at the International AIDS Conference in Durban, the diminutive Nkosi Johnson pleaded with the South African government to start giving the drug azidothymidine (AZT) to pregnant women with HIV so that they would not transmit the virus to their babies. “Don’t be afraid of us — we are all the same,” he told the tearful audience. Johnson, himself HIV-positive from birth, lent a moment of high emotion to a tense week that had been dominated by clashes between scientists, activists and AIDS denialists. The president of South Africa, Thabo Mbeki, had sparked international condemnation when he opened the meeting with a speech that failed to acknowledge HIV as the cause of AIDS.

That week in Durban was a watershed moment for the global AIDS response. As the first international AIDS conference ever held in a developing country, the meeting turned the spotlight on the epidemic in Africa, where the disease was raging worse than anywhere else. In developed countries, antiretroviral drugs (ARVs) had given hope to those living with HIV, but in poor nations, AIDS was still a death sentence for anyone unable to afford the astronomical cost of the medication: about US\$10,000 per person, per year. In the week of the conference alone, an estimated 2,500 South Africans died from AIDS — one-quarter of them children. Johnson died little under a year after giving his speech.

Next week, the International AIDS Conference returns to Durban — but to a radically changed outlook. The government’s AIDS denialists have quietened, and international funding has

poured in. Today, around half of the country's 7 million people with HIV are on ARVs — the biggest such programme in the world. Expanded access to the drugs is largely responsible for a leap in South Africa's average life expectancy at birth, from 53.4 years in 2004 to 62.5 in 2015 (see 'HIV in South Africa'). Mother-to-child transmission has fallen from a high of 30% in the early 2000s to just 1.5%. "It's a miraculous achievement that shows the world what can be done," says Steffanie Strathdee, associate dean of Global Health Sciences at the University of California, San Diego, and one of the headline speakers at this month's conference.

But enormous challenges remain. South Africa still has the largest HIV epidemic in the world, and the rate of new infections remains depressingly high, especially among young women. At the epidemic's epicentre, the province of KwaZulu-Natal, a 15-year-old girl in some communities has an 80% risk of getting HIV in her lifetime. Epidemiologists struggle to fully understand the situation here for the simple reason that those infected are difficult to track. The massive ARV programme is putting immense pressure on the stretched public-health system — and yet the country has just signed up to a major expansion that could double the number of people taking the drugs.

For doctors and medical researchers, the age of ARVs has thrown up new puzzles. They are just beginning to understand how long-term exposure to the drugs and to HIV itself could affect health. Some are exploring what happens to people who reach middle or old age after decades on the treatments. Others are trying to work out whether children who are exposed to HIV and ARVs in the womb might face health problems even if they do not contract the virus.

These results could prove important for the rest of the world, where ARVs could soon be rolled out on an unprecedented scale. The global AIDS-strategy body, UNAIDS, has set a target known as 90-90-90: by 2020, 90% of people living with HIV should know their status; 90% of those diagnosed should be on ARVs; and 90% of those on ARVs should have undetectable levels of the virus. South Africa has been a testing ground — to see whether ARV programmes can be ramped up in developing countries, and to see what happens to the population when they are. "South Africa's success in defeating AIDS is key to the global effort to end AIDS," says Salim Abdool Karim, director of the Durban-based Centre for the AIDS Program of Research in South Africa.

LIFE AFTER HIV

The face of HIV in South Africa today is a young woman called Thembisa Mbhobho. A 15-metre-tall mural of her greets motorists as they turn off the N2 highway to enter Cape Town's Khayelitsha township; it bears the slogan 'There is life beyond HIV'. Mbhobho knows all about that, being one of the millions of South Africans taking daily ARVs to keep the virus at bay. "You can live for more than 50 years if you take your medication correctly," she says. Mbhobho volunteered to be on the mural, which was painted by local artists for World AIDS Day last year, with backing from Médecins sans Frontières (MSF, also known as Doctors without Borders).

"I was diagnosed in 2008. I started taking ARVs in 2014," says the 26-year-old, whose first name means 'promise' in her mother tongue, Xhosa. Today, the virus is so well suppressed in her body that it doesn't show up in blood tests, reducing her likelihood of passing it on. She enjoys it when people recognize her from the mural, or from the television advert that she featured in to encourage people to find out their HIV status. "They motivate me, they say I must keep it up." Her five-year-old son is uninfected, thanks to the drugs Mbhobho was given

while pregnant and in labour. She hopes that one day he'll be a pilot, or a lawyer.

South Africa's road to widespread ARVs has been bumpy. Its first community treatment programme started in 2001 in Khayelitsha, but because of the HIV-denialist ideas subscribed to by Mbeki and other leading politicians, the programme — along with many others in the early days — was labelled as a 'feasibility study'. Some people in the international medical community doubted that ARVs could be effectively administered in Africa. In 2001, the head of the US Agency for International Development, Andrew Natsios, gained notoriety for saying that treatment would not work there because many people in Africa "have not seen a clock or a watch their entire lives", and so would be unable to take their pills on time.

He was wrong. The early studies showed that adherence and treatment outcomes were, in fact, better in African cohorts than in the United States¹. In 2004, when South Africa started to offer free access to ARVs through the public-health system, just under 50,000 citizens received treatment. By 2007, the number was more than 380,000, and today it's well over 3 million.

But the programmes have had different outcomes in South Africa than in high-resource settings in Europe and North America. "In Europe, when ARVs came along, the hospital wards emptied of people who were severely ill," says Gilles van Cutsem, medical coordinator for MSF in South Africa. "When we started our HIV programme in Khayelitsha, the waiting room was full of sick people in wheelbarrows. There is less of that now, but people are still coming in very sick." Despite the information campaigns, and free drugs, many people still wait too long to get tested and treated, he says.

Keeping track of the ARV roll-out has been challenging. The data centre of the Southern Africa International Epidemiological Data-

base to Evaluate AIDS at the University of Cape Town has tracked treatment enrolment and retention since the initiative was established in 2006. But even those simple data have been difficult to collect and interpret, says Morna Cornell, a senior researcher and project manager at the centre. Sometimes the same person turns up in several different records — having moved, perhaps, and started going to a different clinic. Many others are enrolled, but are 'lost to follow-up'. For a long time, the centre's HIV-clinic data showed a low number of AIDS deaths and a high loss to follow-up. But when the data were integrated with the country's death register in the late 2000s, it emerged that more than 30% of those reported as lost were actually dead.

Monitoring should get easier with the planned introduction of unique patient identifiers, says Cornell. She thinks that better data will help South Africa to improve its HIV programmes and assist in future research. Many scientists would like to see much more detailed clinical data collected, so that they can find out what happens when large populations take ARVs in the long term. There is a wealth of research on the health effects of the drugs, but much of it comes from Europe or North America, where patient populations are much smaller, and the medical resources are vastly bigger.

COMPLICATING FACTORS

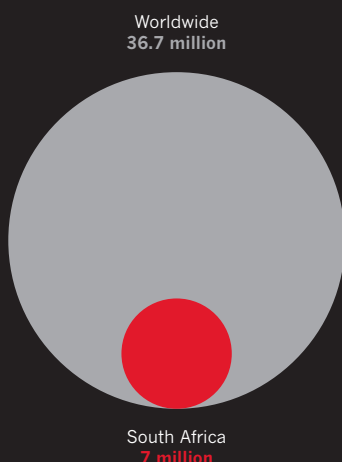
Drug resistance is a looming issue. In rich countries, resistance testing is routinely done on every person diagnosed with HIV to ensure that they are given drugs that will work. But in South Africa, only a few of the available ARV drugs are provided free through the country's health system, and resistance tends to be poorly managed. "People think drug

**"WHEN WE STARTED
OUR HIV PROGRAMME
IN KHAYELITSHA, THE
WAITING ROOM WAS
FULL OF SICK PEOPLE
IN WHEELBARROWS."**

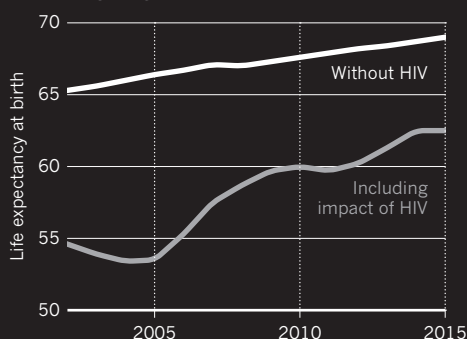
HIV IN SOUTH AFRICA

The country has made huge progress in fighting HIV; access to antiretroviral drugs has soared and life expectancy has risen. But South Africa still has the largest HIV epidemic in the world.

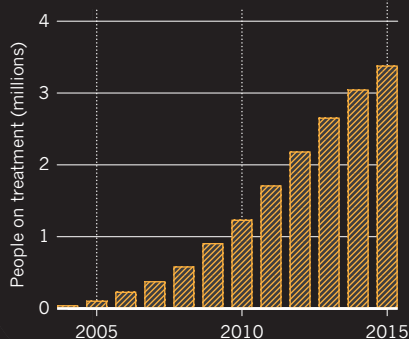
HIV IN 2015



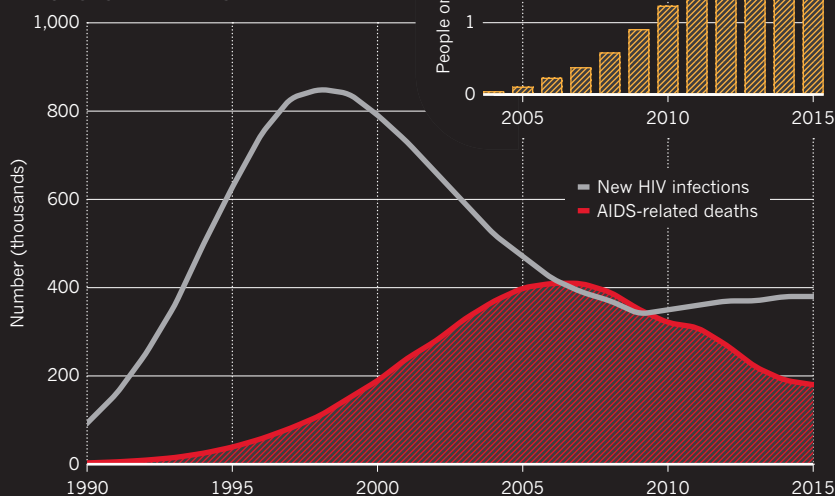
LIFE EXPECTANCY



ANTIRETROVIRAL THERAPY



INFECTIONS AND DEATHS



resistance isn't a problem any more. But the place where resistance is going to emerge again is here. You can't throw so much drug out there and not experience it," says Deenan Pillay, a virologist and director of the Wellcome-Trust-funded Africa Centre for Population Health near Durban.

The Africa Centre holds more than 15 years' worth of longitudinal data, gathered from a community of about 100,000 households in the rural district of uMkhanyakhude, about 2 hours' drive north of Durban, in KwaZulu-Natal. Sure enough, the data show that resistance is a burgeoning problem. Between 2010 and 2012, the proportion of new HIV cases that were already drug resistant rose by 7% — and Pillay suspects that it will grow further. "People sit on failing drug combinations without getting help. That's a big risk," he says.

South Africa's ageing population of people on ARVs is another focus of research. The fraction of HIV-infected people who are older than 50 is predicted to triple in the next 30 years². Some studies have found that

certain cancers are more common in people taking ARVs, and long-term use of the drugs has been linked to increased risk of hypertension, diabetes and obesity³ — although it's difficult to establish whether this is caused by the drugs or by HIV itself. To find out, researchers are now looking at the possible effects of ARVs on metabolism. The life-saving benefits of the drugs outweigh these potential risks, but public-health researchers are nevertheless anxious to know what the future holds for a colossal ARV-taking population.

In Europe and the United States, people ageing with HIV tend unsurprisingly to fare worse health-wise than do those who do not have the virus. But there was a surprise in store for Janet Seeley, a social scientist based at the London School of Hygiene and Tropical Medicine, when she and her colleagues took a close look at data for several hundred people over the age of 50 in Uganda and South Africa. They found that people with HIV who were taking ARVs had better quality of life and were more able to perform daily activities than were their HIV-negative peers⁴.

"My first thought was 'this must be wrong,' but it's logical really," she says. Care for people with HIV isn't always great, but at least they see a health practitioner on a regular basis. In South Africa, older people who received HIV therapy were more likely to be on treatment for other chronic conditions than were people in the comparison group who were HIV-negative.

Researchers are also exploring what happens to children who are born to mothers with the virus but are not infected themselves. Clinicians have found that these 'uninfected exposed' children tend to have worse health — such as lower birth weight and bone density — than do those born to mothers without the virus. But the evidence is conflicting: a systematic review published earlier this year⁵ found some studies indicating that uninfected exposed children are at increased risk of death, hospitalization and under-nutrition, and others that saw no such evidence.

Even if the effects are marginal, the fact that around 30% of children born in South Africa fall into this category is cause to take the issue

very seriously. Mark Cotton, head of infectious diseases at Tygerberg Children's Hospital in Cape Town, is convinced from what he has observed in the clinic that these children fare worse than others. The challenge lies in working out why. "There are theories that HIV-positive women might be sicker, or poorer, or that there might be more TB in these homes. But we think it could go beyond that," he says.

Clive Gray, an immunologist at the University of Cape Town, is investigating the possible effects of *in utero* HIV and ARV exposure on the developing fetus and on an uninfected child's long-term health. With funding from the Canadian Institutes of Health Research, he and some colleagues are investigating 500 mother-baby pairs — including mothers both with and without the virus — in Nigeria and South Africa. The study is not complete, but Gray and his colleagues have already spotted that HIV-negative children born to infected mothers have impaired immunity in their first year of life.

As to whether HIV or the ARVs are the cause, Gray says, that is

SOURCES: UNAIDS AND STATISTICS SOUTH AFRICA

difficult to tease out. And he's not sure it's the most important question to answer. "These children are sick, and whether that is because they are exposed to HIV or to the drug, or to the combination, it doesn't matter from a public-health point of view." Ultimately, the work might point to alternative ways to handle HIV in pregnant women. And, Gray says, the research might also help to explain why women with the virus often go into preterm labour — another confounding factor, because it is known to negatively affect child development.

BREAK THE CYCLE

But perhaps the most pressing question is how South Africa's massive rollout of ARVs will affect infection rates. People who have a suppressed viral load are less likely to transmit the virus to others — a fact that led the World Health Organization (WHO) to recommend last year that all people who test positive for HIV immediately go on ARVs, rather than waiting until their CD4 count — an indicator of disease progression — fell below a certain level. In May, South Africa's health minister, Aaron Motsoaledi, announced that the country would adopt these guidelines from September, a strategy that could double the number who receive ARVs. Motsoaledi also said that his department

would supply prophylactic treatment to sex workers, who have a very high risk of contracting HIV.

Even patchy ARV coverage can slash infection rates. A study by the Africa Centre found that someone living in a community where 30–40% of HIV-infected people are on treatment is 38% less likely to get infected than is someone living in a community where fewer than 10% are on treatment⁶.

In practice, however, ARV therapy is unlikely to ramp up quickly. A major difficulty lies in getting people tested in the first place, and another is getting those who test positive to accept treatment — because many of them will not feel sick — and to stay on it.

To reach those people, the country has to rethink its HIV programmes, van Cutsem says. Research by MSF has found that allowing people who manage their HIV well to attend fewer clinics helps to take the strain off the facilities and health workers. He also supports the introduction of self-testing kits.

These can already be bought over the counter in South African pharmacies, but are not promoted in the public-health sector. "There was an initial concern in South Africa that it would lead to self-harm or domestic violence. But frankly, there is no evidence saying the risk would be bigger than if you test with a counsellor," van Cutsem says.

There is no shortage of scientists to study the evolving epidemic. Since South Africa hosted the International AIDS Conference in 2000, many world-class research centres have emerged to develop vaccines and therapies that might, one day, supersede ARVs. And later this year, the country will start a large trial of an HIV vaccine candidate that has shown promising results in Thailand⁷. It would be the first major vaccine trial in years and one that, if successful, could make a big difference, especially to women with the highest risk of infection.

For many of the researchers returning to Durban next week, the trip will remind them of how far they've come. "Before ARVs, HIV was the same around the world, in that everybody died who got it," says Carlos del Rio, a global-health researcher at Emory University School of Medicine in Atlanta, Georgia, and one of the keynote speakers in Durban this year. "But after 1996, people stopped dying in the developed world. It was that meeting in Durban that made us believe that the impossible would be possible. Now people live in South Africa with HIV who would otherwise have died."

But the conference also makes del Rio reflect on how far there is to go. A huge number of African people are reaching reproductive age, and encouraging them to understand and manage their HIV risk will be crucial to sustaining the continent's momentum. Some would not have seen the epidemic at its chilling peak, and could become complacent about the need to protect themselves and others. "We run a massive risk that if we fail with that generation, the epidemic could get even worse."

"Are we at the end of AIDS?" says del Rio. "I don't think so. I think we're at the beginning of the end of AIDS. But not keeping up investment, not keeping our eyes on the ball, could be potentially devastating." ■

Linda Nordling is a reporter based in Cape Town, South Africa.

1. Mills, E. J. *et al. J. Am. Med. Assoc.* **296**, 679–690 (2006).
2. Hontelez, J. A. C. *et al. AIDS* **25**, 1665–1667 (2011).
3. Julius, H. *et al. Curr. HIV Res.* **9**, 247–252 (2011).
4. Nyirenda, M. *et al. BMC Publ. Health* **12**, 259 (2012).
5. le Roux, S. M., Abrams, E. J., Nguyen, K. & Myer, L. *Trop. Med. Int. Health* **21**, 829–845 (2016).
6. Tanser, F. *et al. Science* **339**, 966–971 (2013).
7. Rerks-Ngarm, S. *et al. N. Engl. J. Med.* **361**, 2209–2220 (2009).

"IT WAS THAT MEETING IN DURBAN THAT MADE US BELIEVE THAT THE IMPOSSIBLE WOULD BE POSSIBLE."



Thembisa Mbhobho, who is HIV-positive, is depicted on a mural near Cape Town's Khayelitsha township.



THE FORGOTTEN CONTINENT

Fossil finds in China are challenging ideas about the evolution of modern humans and our closest relatives.

BY JANE QIU

On the outskirts of Beijing, a small limestone mountain named Dragon Bone Hill rises above the surrounding sprawl. Along the northern side, a path leads up to some fenced-off caves that draw 150,000 visitors each year, from schoolchildren to grey-haired pensioners. It was here, in 1929, that researchers discovered a nearly complete ancient skull that they determined was roughly half a million years old. Dubbed Peking Man, it was among the earliest human remains ever uncovered, and it helped to convince many researchers that humanity first evolved in Asia.

Since then, the central importance of Peking Man has faded. Although modern dating methods put the fossil even earlier — at up to 780,000 years old — the specimen has been eclipsed by discoveries in Africa that have yielded much older remains of ancient human relatives. Such finds have cemented Africa's status as the cradle of humanity — the place from which modern humans and their predecessors

spread around the globe — and relegated Asia to a kind of evolutionary cul-de-sac.

But the tale of Peking Man has haunted generations of Chinese researchers, who have struggled to understand its relationship to modern humans. "It's a story without an ending," says Wu Xinzhi, a palaeontologist at the Chinese Academy of Sciences' Institute of Vertebrate Paleontology and Paleoanthropology (IVPP) in Beijing. They wonder whether the descendants of Peking Man and fellow members of the species *Homo erectus* died out or evolved into a more modern species, and whether they contributed to the gene pool of China today.

Keen to get to the bottom of its people's ancestry, China has in the past decade stepped up its efforts to uncover evidence of early humans across the country. It is reanalysing old fossil finds and pouring tens of millions of dollars a year into excavations. And the

government is setting up a US\$1.1-million laboratory at the IVPP to extract and sequence ancient DNA.

The investment comes at a time when palaeo-anthropologists across the globe are starting to pay more attention to Asian fossils and how they relate to other early hominins — creatures that are more closely related to humans than to chimps. Finds in China and other parts of Asia have made it clear that a dazzling variety of *Homo* species once roamed the continent. And they are challenging conventional ideas about the evolutionary history of humanity.

"Many Western scientists tend to see Asian fossils and artefacts through the prism of what was happening in Africa and Europe," says Wu. Those other continents have historically drawn more attention in studies of human evolution because of the antiquity of fossil finds there, and because they are closer to major palaeoanthropology research institutions, he says. "But it's increasingly clear that many Asian

DEAGOSTINI/GETTY

The reconstructed skull of Peking Man, the fossil that launched discussions of human origins in China.

materials cannot fit into the traditional narrative of human evolution.”

Chris Stringer, a palaeoanthropologist at the Natural History Museum in London, agrees. “Asia has been a forgotten continent,” he says. “Its role in human evolution may have been largely under-appreciated.”

EVOLVING STORY

In its typical form, the story of *Homo sapiens* starts in Africa. The exact details vary from one telling to another, but the key characters and events generally remain the same. And the title is always ‘Out of Africa’.

In this standard view of human evolution, *H. erectus* first evolved there more than 2 million years ago (see ‘Two routes for human evolution’). Then, some time before 600,000 years ago, it gave rise to a new species: *Homo heidelbergensis*, the oldest remains of which have been found in Ethiopia. About 400,000 years ago, some members of *H. heidelbergensis* left Africa and split into two branches: one ventured into the Middle East and Europe, where it evolved into Neanderthals; the other went east, where members became Denisovans — a group first discovered in Siberia in 2010. The remaining population of *H. heidelbergensis* in Africa eventually evolved into our own species, *H. sapiens*, about 200,000 years ago. Then these early humans expanded their range to Eurasia 60,000 years ago, where they replaced local hominins with a minuscule amount of interbreeding.

A hallmark of *H. heidelbergensis* — the potential common ancestor of Neanderthals, Denisovans and modern humans — is that individuals have a mixture of primitive and modern features. Like more archaic lineages, *H. heidelbergensis* has a massive brow ridge and no chin. But it also resembles *H. sapiens*, with its smaller teeth and bigger braincase. Most researchers have viewed *H. heidelbergensis* — or something similar — as a transitional form between *H. erectus* and *H. sapiens*.

Unfortunately, fossil evidence from this period, the dawn of the human race, is scarce and often ambiguous. It is the least understood episode in human evolution, says Russell Ciochon, a palaeoanthropologist at the University of Iowa in Iowa City. “But it’s central to our understanding of humanity’s ultimate origin.”

The tale is further muddled by Chinese fossils analysed over the past four decades, which cast doubt over the linear progression from African *H. erectus* to modern humans. They show that, between roughly 900,000 and 125,000 years ago, east Asia was teeming with hominins endowed with features that would place them somewhere between *H. erectus* and *H. sapiens*, says Wu.

“Those fossils are a big mystery,” says Ciochon. “They clearly represent more advanced species than *H. erectus*, but nobody

knows what they are because they don’t seem to fit into any categories we know.”

The fossils’ transitional characteristics have prompted researchers such as Stringer to lump them with *H. heidelbergensis*. Because the oldest of these forms, two skulls uncovered in Yunxian in Hubei province, date back 900,000 years^{1,2}, Stringer even suggests that *H. heidelbergensis* might have originated in Asia and then spread to other continents.

But many researchers, including most Chinese palaeontologists, contend that the materials from China are different from European and African *H. heidelbergensis* fossils, despite some apparent similarities. One nearly complete skull unearthed at Dali in Shaanxi province and dated to 250,000 years ago, has a bigger braincase, a shorter face and a lower cheekbone than most *H. heidelbergensis* specimens³, suggesting that the species was more advanced.

THE EVOLUTIONARY TALE IN ASIA IS MUCH MORE INTERESTING THAN PEOPLE REALIZED.

Such transitional forms persisted for hundreds of thousands of years in China, until species appeared with such modern traits that some researchers have classified them as *H. sapiens*. One of the most recent of these is represented by two teeth and a lower jawbone, dating to about 100,000 years ago, unearthed in 2007 by IVPP palaeoanthropologist Liu Wu and his colleagues⁴. Discovered in Zhirendong, a cave in Guangxi province, the jaw has a classic modern-human appearance, but retains some archaic features of Peking Man, such as a more robust build and a less-protruding chin.

Most Chinese palaeontologists — and a few ardent supporters from the West — think that the transitional fossils are evidence that Peking Man was an ancestor of modern Asian people. In this model, known as multiregionalism or continuity with hybridization, hominins descended from *H. erectus* in Asia interbred with incoming groups from Africa and other parts of Eurasia, and their progeny gave rise to the ancestors of modern east Asians, says Wu.

Support for this idea also comes from artefacts in China. In Europe and Africa, stone tools changed markedly over time, but hominins in China used the same type of simple stone instruments from about 1.7 million years ago to 10,000 years ago. According to Gao Xing, an archaeologist at the IVPP, this suggests that local hominins evolved continuously, with little influence from outside populations.

POLITICS AT PLAY?

Some Western researchers suggest that there is a hint of nationalism in Chinese palaeontologists’ support for continuity. “The Chinese — they do not accept the idea that *H. sapiens* evolved in Africa,” says one researcher. “They

want everything to come from China.”

Chinese researchers reject such allegations. “This has nothing to do with nationalism,” says Wu. It’s all about the evidence — the transitional fossils and archaeological artefacts, he says. “Everything points to continuous evolution in China from *H. erectus* to modern human.”

But the continuity-with-hybridization model is countered by overwhelming genetic data that point to Africa as the wellspring of modern humans. Studies of Chinese populations show that 97.4% of their genetic make-up is from ancestral modern humans from Africa, with the rest coming from extinct forms such as Neanderthals and Denisovans⁵. “If there had been significant contributions from Chinese *H. erectus*, they would show up in the genetic data,” says Li Hui, a population geneticist at Fudan University in Shanghai. Wu counters that the genetic contribution from archaic hominins in China could have been missed because no

DNA has yet been recovered from them.

Many researchers say that there are ways to explain the existing Asian fossils without resorting to continuity with hybridization. The Zhirendong hominins, for instance, could represent an exodus of early modern humans from Africa between 120,000 and 80,000 years ago. Instead of remaining in the Levant in the Middle East, as was thought previously, these people could have expanded into east Asia, says Michael Petraglia, an archaeologist at the University of Oxford, UK.

Other evidence backs up this hypothesis: excavations at a cave in Daoxian in China’s Hunan province have yielded 47 fossil teeth so modern-looking that they could have come from the mouths of people today. But the fossils are at least 80,000 years old, and perhaps 120,000 years old, Liu and his colleagues reported last year⁶. “Those early migrants may have interbred with archaic populations along the way or in Asia, which could explain Zhirendong people’s primitive traits,” says Petraglia.

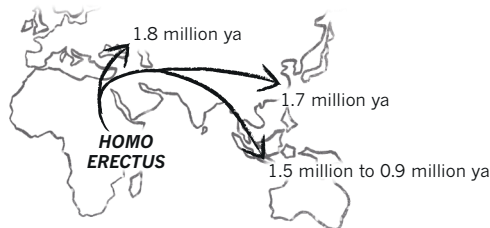
Another possibility is that some of the Chinese fossils, including the Dali skull, represent the mysterious Denisovans, a species identified from Siberian fossils that are more than 40,000 years old. Palaeontologists don’t know what the Denisovans looked like, but studies of DNA recovered from their teeth and bones indicate that this ancient population contributed to the genomes of modern humans, especially Australian Aborigines, Papua New Guineans and Polynesians — suggesting that Denisovans might have roamed Asia.

María Martínón-Torres, a palaeoanthropologist at University College London, is among those who proposed that some of the Chinese

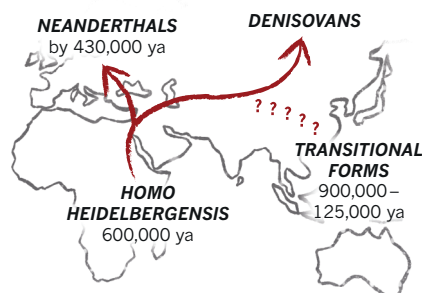
TWO ROUTES FOR HUMAN EVOLUTION

AFRICAN ORIGIN

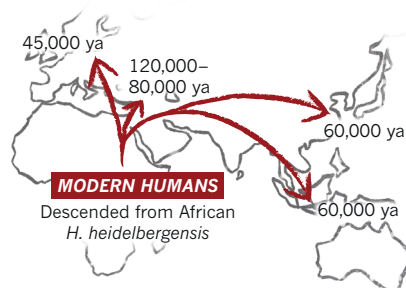
Homo erectus evolved in Africa and had dispersed into Asia by 1.8 million years ago (ya).



OUT OF AFRICA

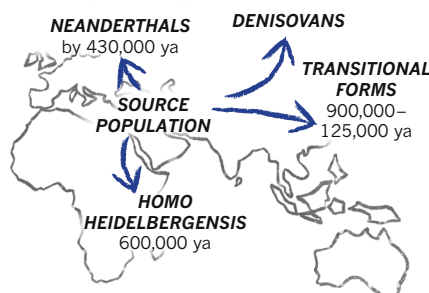


Homo heidelbergensis or other transitional forms evolved into Neanderthals and Denisovans. Hominins in China with mixtures of archaic and modern features may have derived from *H. heidelbergensis*.

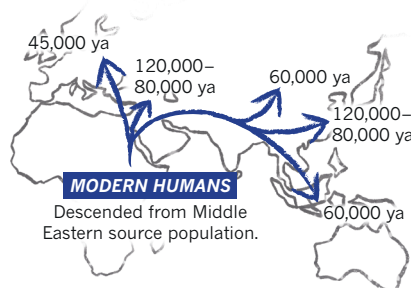


Modern humans emerged in Africa about 200,000 years ago and reached the Middle East by 120,000–80,000 years ago. Later waves of modern humans spread through Europe and Asia.

ALTERNATIVE MODEL



Homo erectus descendants in the Middle East give rise to various hominin groups in Europe and east Asia, as well as to *Homo heidelbergensis* in Africa.



Modern humans evolve in Africa from *H. heidelbergensis* or another hominin derived from the Middle East, and they disperse to Eurasia in multiple waves.

hominins were Denisovans. She worked with IVPP researchers on an analysis⁷, published last year, of a fossil assemblage uncovered at Xujiayao in Hebei province — including partial jaws and nine teeth dated to 125,000–100,000 years ago. The molar teeth are massive, with very robust roots and complex grooves, reminiscent of those from Denisovans, she says.

A third idea is even more radical. It emerged when Martínón-Torres and her colleagues compared more than 5,000 fossil teeth from around the world: the team found that Eurasian specimens are more similar to each other than to African ones⁸. That work and more recent interpretations of fossil skulls suggest that Eurasian hominins evolved separately from African ones for a long stretch of time. The researchers propose that the first hominins that left Africa 1.8 million years ago were the eventual source

of modern humans. Their descendants mostly settled in the Middle East, where the climate was favourable, and then produced waves of transitional hominins that spread elsewhere. One Eurasian group went to Indonesia, another gave rise to Neanderthals and Denisovans, and a third ventured back into Africa and evolved into *H. sapiens*, which later spread throughout the world. In this model, modern humans evolved in Africa, but their immediate ancestor originated in the Middle East.

Not everybody is convinced. “Fossil interpretations are notoriously problematic,” says Svante Pääbo, a palaeogeneticist at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. But DNA from Eurasian fossils dating to the start of the human race could help to reveal which story — or combination — is correct. China is now

making a push in that direction. Qiaomei Fu, a palaeogeneticist who did her PhD with Pääbo, returned home last year to establish a lab to extract and sequence ancient DNA at the IVPP. One of her immediate goals is to see whether some of the Chinese fossils belong to the mysterious Denisovan group. The prominent molar teeth from Xujiayao will be an early target. “I think we have a prime suspect here,” she says.

FUZZY PICTURE

Despite the different interpretations of the Chinese fossil record, everybody agrees that the evolutionary tale in Asia is much more interesting than people appreciated before. But the details remain fuzzy, because so few researchers have excavated in Asia.

When they have, the results have been startling. In 2003, a dig on Flores island in Indonesia turned up a diminutive hominin⁹, which researchers named *Homo floresiensis* and dubbed the hobbit. With its odd assortment of features, the creature still provokes debate about whether it is a dwarfed form of *H. erectus* or some more primitive lineage that made it all the way from Africa to southeast Asia and lived until as recently as 60,000 years ago. Last month, more surprises emerged from Flores, where researchers found the remains of a hobbit-like hominin in rocks about 700,000 years old¹⁰.

Recovering more fossils from all parts of Asia will clearly help to fill in the gaps. Many palaeoanthropologists also call for better access to existing materials. Most Chinese fossils — including some of the finest specimens, such as the Yunxian and Dali skulls — are accessible only to a handful of Chinese palaeontologists and their collaborators. “To make them available for general studies, with replicas or CT scans, would be fantastic,” says Stringer. Moreover, fossil sites should be dated much more rigorously, preferably by multiple methods, researchers say.

But all agree that Asia — the largest continent on Earth — has a lot more to offer in terms of unravelling the human story. “The centre of gravity,” says Petraglia, “is shifting eastward.” ■

Jane Qiu is a writer in Beijing.

- Li, T. & Etler, D. A. *Nature* **357**, 404–407 (1992).
- Vialet, A. et al. *Comptes Rendus Palevol* **9**, 331–339 (2010).
- Wu, X. & Athreya, S. *Am. J. Phys. Anthropol.* **150**, 141–157 (2013).
- Liu, W. et al. *Proc. Natl Acad. Sci. USA* **107**, 19201–19206 (2010).
- Hu, Y. et al. Preprint at <http://arxiv.org/abs/1404.7766> (2014).
- Liu, W. et al. *Nature* **526**, 696–699 (2015).
- Xing, S., Martínón-Torres, M., Bermúdez de Castro, J. M., Wu, X. & Liu, W. *Am. J. Phys. Anthropol.* **156**, 224–240 (2015).
- Martínón-Torres, M. et al. *Proc. Natl Acad. Sci. USA* **104**, 13279–13282 (2007).
- Brown, P. et al. *Nature* **431**, 1055–1061 (2004).
- van den Bergh, G. D. et al. *Nature* **534**, 245–248 (2016).

COMMENT



CONSERVATION Seas need both marine protected areas and fisheries management **p.224**

BOOKS Birds, beasts and the brain: holiday reads from *Nature* reviewers **p.228**

DEVELOPMENT African nations must invest to stem brain drain **p.231**

OBITUARY Jerome Bruner, cognitive psychology pioneer, remembered **p.232**

LALO DE ALMEIDA/CONTRASTO/EYEVINE



A truck full of sugar cane crosses farmland in Brazil, a leader in bio-based ethanol production from the crop.

Five cornerstones of a global bioeconomy

Beate El-Chichakli and colleagues outline principles for coordinating bio-based industries to achieve many of the sustainable development goals.

More than 40 nations¹ are proposing to boost their ‘bioeconomy’ — the part of the economy based in biology and the biosciences. Around US\$2 trillion of products in agriculture and forestry, food, bioenergy, biotechnology and green chemistry were exported worldwide in 2014, amounting² to 13% of world trade, up from 10% in 2007. These sectors are central to at least half of the UN Sustainable Development Goals (SDGs), from food security to ensuring energy access and health. But conflicting national priorities make it hard to align bioeconomy policies to meet the SDGs on a global scale.

Bioscience leaders such as the European

Union, Japan and the United States see expanding the bioeconomy as a means of reindustrializing and creating wealth (see ‘Bioeconomy breakdown’). Emerging industrial economies such as China and India see biotechnology as a nascent field of innovation in which they can quickly compete. Brazil, South Africa and Malaysia are investing to add value to their vast biological resources. Ecological sustainability is a prime concern in rich and industrializing countries; inclusive rural development and equitable sharing of resources is central in developing countries.

Knock-on effects of decisions made in one place may be felt elsewhere. The EU’s plan to ecologically certify its bio-based products will

impede producers in poorer countries who lack testing infrastructure. In India and Brazil, strict bio-piracy regulations, which aim to protect biodiversity and traditional knowledge, are failing to benefit local populations and are stalling international research in plant biodiversity.

Without agreed global priorities and assessment methods, it is hard to take into account such indirect effects and trade-offs. Global investments in information collection such as satellite tracking of biomass call for joint action. Nations will fund DNA barcoding of rare rainforest hardwoods or use genetic stock identification to manage fisheries only when such technologies are ►

INNOVATIONS IN THE BIOECONOMY

Scoring Sustainable Development Goals

SDG 2: END HUNGER Food security is the top priority⁵. More-efficient animal production and meat substitutes are needed. Chicken is more sustainable than beef, owing to lower greenhouse-gas emissions and water needs. Genomic technologies will need to be applied to more foods, as they have been to dairy cattle, chicken, salmon, tilapia, rice and banana. Farmed seafood production must be boosted and will require new vaccines and molecular diagnostics to reduce antibiotic use, as well as sources of protein-rich feed.

SDG 3: ENSURE HEALTHY LIVES Sustainable medicines, such as biopharmaceuticals, and microbiome-based approaches are needed for infectious diseases such as malaria and epidemics including diabetes and obesity. For example, production of semi-synthetic artemisinin from microbially sourced artemisinic acid is an early success story for combining metabolic engineering and synthetic biology in the commercial production of drugs against malaria⁶.

SDG 6: WATER AND SANITATION FOR ALL In developing countries, 90% of sewage and 70% of industrial wastes are discharged without treatment. Advances in biological wastewater treatment, including phosphorus removal and nitrification, hold potential if implemented more widely. Small, modular systems should be spread to remote communities, and large, intensive plants can cater for city-sized populations.

SDG 7: ENERGY FOR ALL Most developing countries have unreliable energy systems. Burning wood or manure leads to health problems, premature deaths and deforestation. Decentralized, modern solutions that combine bioenergy with other renewables are needed. For example, an Indian social enterprise has implemented dairy and biogas production and local mini-grids electrified by biogas from waste or by eco-briquettes.

SDG 8 & 9: SUSTAINABLE ECONOMIC GROWTH Combining rural regeneration with reindustrialization offers increased sustainability and inclusiveness. Brazil, a leader in bio-based ethanol production with around 300 operating sugar-cane or ethanol mills, is commercializing production of cellulosic ethanol. In Finland, a large biorefinery currently under

construction will need only 200 people to run it but creates another 2,500 jobs across the value chain for growing, harvesting and transporting biomass.

SDG 11: SUSTAINABLE CITIES Biological principles — such as metabolisms, ecosystems and cycles — can be applied to help cities to function sustainably¹. Local production and recycling systems minimize emissions and waste. Renewable resources, cultivation methods and biotechnology can close material and energy cycles and loops. For example, Edmonton in Canada is aiming to recover 90% of its organic waste and convert domestic waste into biofuels.

SDG 12: SUSTAINABLE CONSUMPTION Bio-based materials and chemicals are increasingly deployed in industries such as plastics processing, consumer goods, construction, pharmaceuticals and medical technology. Switching from fossil to bio-based materials would make a big difference in the chemicals sector, which has the third-largest emissions in industry, after steel and cement. Using enzymes in detergents has enabled consumers to significantly reduce washing temperature, for example.

SDG 13: COMBAT CLIMATE CHANGE Bio-based industries are active in carbon storage and mitigating climate change. Biotechnology companies are collaborating with heavy industry to make carbon dioxide emissions into bio-based chemicals and biofuels⁷.

SDG 14: OCEANS, SEAS AND MARINE RESOURCES Illegal, unreported and unregulated fishing remains a major threat to marine ecosystems⁸. The need for traceability is urgent — one-third of the world's fisheries catch from 1950–2002 lacked species identification. The use of DNA barcodes in a global database — the Barcode of Life Data System — could be expanded to address traceability and fish fraud as well as yielding information on migration and dispersal.

SDG 15: TERRESTRIAL ECOSYSTEM With limited land area, agriculture must be intensified. And farming must decouple from the fossil-fuel industry. Advanced breeding technologies can avoid soil exhaustion and degradation. Self-fertilizing versions of food staples such as maize (corn), wheat, barley and rice, and fertilizing soils by microbial communities, could become feasible through bioengineering in a decade.

▶ internationally used and enforced.

In November 2015, more than 700 experts from around 80 countries met in the first Global Bioeconomy Summit in Berlin. We, as members of the International Advisory Committee on the Bioeconomy (comprising 37 experts from around the globe who shaped the summit) outline the principles that were agreed and the steps needed to advance them, as well as illustrating how these can be applied to individual SDGs (see 'Scoring Sustainable Development Goals').

DIFFICULT BALANCE

A global bioeconomy must rebuild natural capital and improve the quality of life for a growing world population. It should balance managing common goods, such as air, water and soil, with the economic expectations of people. Three types of innovation will be needed: technological (such as systems to reduce emissions), organizational (changes in institutional behaviour) and social (such as job creation). For example, new sorts of sustainable building materials based on wood or lignin (a compound found in many plant cell walls) will need to be integrated into building codes. Also needed will be citizen-science evaluations of new houses, local wood-recycling and construction efforts. Sustainable food systems will require advances in plant breeding, food products, and farming and cultivation techniques, as well as steps to optimize shelf-life and food distribution, and social initiatives such as the revival of traditional crops, food-sharing platforms and low-meat diets.

Inclusiveness and knowledge transfer are important. For example, biotechnology infrastructure and skilled employees are found mostly in high-income countries, whereas local biological know-how and reuse culture are strong in developing countries. Supply chains should create local jobs, with manufacturing close to the raw-material base. For example, bioenergy mini-grids or bricks-from-waste production plants increase local energy access and jobs in rural India.

Regulatory frameworks for intellectual property, the access to and use of genetic resources, biosafety and the ethics of biosciences and industrial standards all need to be reviewed. Globally agreed standards on the measurement and definition of bio-based products — such as the carbon footprint and sustainability of bioplastics — are needed. A certifying and testing body must be independent and international to establish public confidence and enable countries that lack capacities to benefit from the results.

FIVE CORNERSTONES

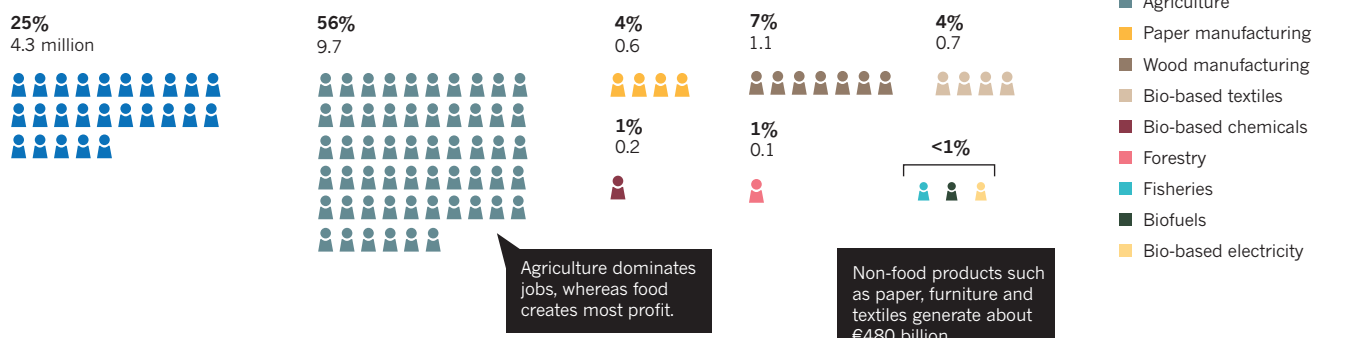
Unifying principles for a global bioeconomy need to be put in place by international policy bodies, multilateral trade negotiators

SOURCE: REF. 9 (EU)

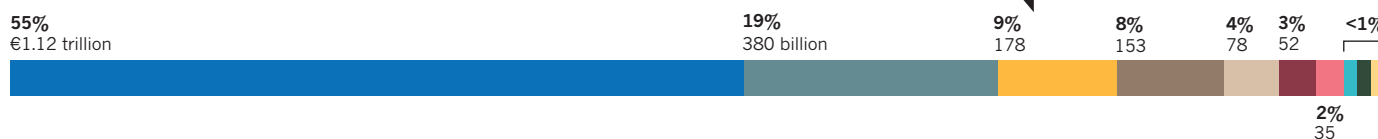
BIOECONOMY BREAKDOWN

The European Union's biology-based industries account for 17 million jobs, or 8.5% of the region's workforce (A) and generate more than €2 trillion (US\$2.2 trillion) annually (B). Equivalent statistics are unavailable elsewhere, but US industries producing bio-based products (non-food) account for about 4 million jobs and US\$370 billion. India's bioeconomy surpassed US\$4 billion in 2013. And Brazil's sugar-cane industry accounted for 2% of its gross domestic product and 4.5 million jobs in 2012.

A EU's bioeconomy employment: 17 million jobs



B EU's bioeconomy turnover: €2 trillion



and the corporate sector.

First, international collaborations between governments and public and private researchers are essential for optimizing resource use and sharing knowledge. For instance, international initiatives that bring together bioscience and IT are needed in sustainable intensification and global soil mapping, to agree common protocols. The Food and Agriculture Organization of the United Nations (FAO) established a Sustainable Bioeconomy Working Group in 2016 and could develop such a forum.

Second, ways to measure the bioeconomy's development and its contributions to the SDGs need to be found. Priority targets, such as food security, and assessment criteria need to be agreed internationally, led by global organizations such as the UN and its subsidiaries. National monitoring systems should include the international dimension so that a country could examine how its practices might affect others. This will entail efforts to make related data openly accessible.

Third, bioeconomy initiatives need to be linked more closely with multilateral policy processes and intergovernmental discussions, particularly the SDG 2030 agenda and follow-ups to the Paris climate and Aichi biodiversity agreements. A UN body on bioeconomy needs to be set up to handle the coordination. Carbon pricing, the costing of other negative impacts (including indirect costs such as air pollution and climate change) and the removal of fossil-fuel subsidies are necessities for meeting the SDGs. Fossil energy received a staggering \$5.3 trillion, or 6.5% of global gross domestic product, in post-tax subsidies in 2015³.

Fourth, educators should collaborate

internationally to define the knowledge, skills and competencies required for developing a bioeconomy that enhances the sustainable use of bio-based materials in manufacturing and in consumer products. This will require an interdisciplinary approach that emphasizes systems thinking, strategic planning and evaluating environmental, social and economic performance, as well as an understanding of technologies and local specifics. Governments should build international teaching and learning exchanges into research programmes. Open learning platforms would allow the sharing of curricula and training content.

Fifth, research-and-development support programmes are needed to encourage global collaborations in a few breakthrough projects. For example, 300 experts concluded⁴ that collaboration in bioeconomy research would be most needed in the following areas: new food systems, bio-principled cities, sustainable aquaculture, biorefineries, artificial photosynthesis, consumer and citizen participation and global governance.

Discussions on these five cornerstones should begin now, so that structures can be in place before the next Global Bioeconomy Summit in 2017. ■

Beate El-Chichakli is head of the Office of the Bioeconomy Council of the German Federal Government, Berlin, Germany.

Joachim von Braun is director of the Center for Development Research and Professor for Economic and Technological Change at the University of Bonn, Germany, and co-chair of the Bioeconomy Council of the German Federal Government. **Christine Lang** is

professor of microbiology and molecular genetics at the Technical University of Berlin, Germany, and co-chair of the Bioeconomy Council of the German Federal Government. **Daniel Barben** is professor and director of the Institute of Science, Technology and Society Studies at the Alpen-Adria-Universität Klagenfurt, Austria, and a member of the Bioeconomy Council of the German Federal Government. **Jim Philp** is a policy analyst at the Organisation for Economic Co-operation and Development, Paris, France.
e-mail: james.philp@oecd.org

1. Communiqué of the Global Bioeconomy Summit. *Making Bioeconomy Work for Sustainable Development* (2015); available at <http://go.nature.com/293zhq2>
2. Intesa SanPaolo. *La bioeconomia in Europa: 2° Rapporto* (2015); available at <http://go.nature.com/29as14n> (in Italian).
3. Coady, D., Parry, I., Sears, L. & Shang, B. *How Large Are Global Energy Subsidies?* (International Monetary Fund, 2015).
4. Bioökonomierat. *Global Visions for the Bioeconomy: An international Delphi-Study* (Office of the German Bioeconomy Council, 2015); available at <http://go.nature.com/29rz1fx>
5. von Braun, J. in *The Fight Against Hunger and Malnutrition* (ed. Sahn, D.) 240–262 (Oxford Univ. Press, 2015).
6. US National Academy of Sciences. *Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals* (National Academies Press, 2015).
7. Peplow, M. *Nature Biotechnol.* **33**, 1123–1125 (2015).
8. Food and Agriculture Organization of the United Nations. *The State of World Fisheries and Aquaculture 2014 Highlights* (FAO, 2014).
9. Ronzon, T., Santini, F. & M'Barek, R. *The Bioeconomy in the European Union in Numbers* (2015); available at <http://go.nature.com/291rc3l>

The views expressed are those of the authors and not necessarily those of the OECD or of the governments of its member countries.



Fishing off the east African coast near Vamizi Island, only one side of which is a marine protected area.

Marine biodiversity needs more than protection

To sustain the seas, advocates of marine protected areas and those in fisheries management must work together, not at cross purposes, urges **Ray Hilborn**.

On 1 September, government leaders, directors of non-governmental organizations (NGOs) and others will meet in Hawaii at the International Union for Conservation of Nature's World Conservation Congress to discuss environmental and development challenges. Twenty-three NGOs, including the Pew Charitable Trusts and the Natural Resources Defense Council, are calling on the IUCN to make 30% of the world's coastal and marine areas fully protected from fishing and other forms of exploitation by 2030.

If this target were achieved, the abundance of exploited species in the areas that are closed off would undoubtedly increase¹. It is not clear, however, whether the same would be true for marine biodiversity overall.

There are currently two very different views on the effectiveness of zones where

fishing is either banned outright or tightly restricted. Many conservationists see the establishment of these marine protected areas (MPAs) as the only way to protect biodiversity. Others — me included — argue that the protection of biodiversity at sea can include recreational and industrial fishing and other uses of ocean resources. In fact, we think that closing waters to some kinds of fishing gear and restricting the catch of named species can offer much more protection than cordoning off even 30% of an area. We are concerned that MPAs may simply shift fishing pressure elsewhere².

Opinions are so divided that the conservation expertise of fisheries managers is being left out of national and international drives to protect ocean resources. Likewise, the suite of threats to biodiversity besides fishing, such as from oil exploration,

sea-bed mining and ocean acidification, are not being addressed in standard fisheries management.

The seas face myriad problems — climate change, development and the nutritional and other needs of a growing human population. To tackle them, conservationists and those involved in fisheries management must work together and answer to the same governing bodies.

RISE OF PROTECTION

Calls for MPAs began in earnest during the 1990s, when overfishing was common in most of the developed world and collapses of fish stocks repeatedly made headlines. In the early 2000s, ecologists often assumed that biodiversity could flourish only inside protected areas. One group proposed in 2002, for example, that 40% of the ocean be

made reserves, on the assumption that the replenishment of fish populations through reproduction could not happen outside them³.

Most ecologists and conservationists now accept — in theory — that even if as much as 20% of a region were cordoned off from fishing, most of that area's biodiversity would exist outside the protected zones as long as effective fisheries management was in place. Yet the dominance of MPAs in conservation policy has, if anything, increased since the 2000s.

In the past decade especially, numerous environmental NGOs and conservation-funding groups have taken up MPAs as their preferred tool for ocean protection. Together, the conservation group WWF, Greenpeace and other NGOs have spent hundreds of millions of dollars over the past ten years lobbying for MPAs around the world. One effect of this was US President Barack Obama adding just over 1 million square kilometres (an area roughly twice the size of Texas) to the US Pacific territories national monument in 2014. Another has been President James Michel of the Seychelles promising to make 412,000 km² of the Indian Ocean surrounding the islands a totally protected MPA.

MPAs also dominate the scientific literature on marine conservation. Researchers documenting the effects of MPAs on biodiversity, in my view, ignore or underappreciate the benefits of fisheries management. Jane Lubchenco and Kirsten Grorud-Colvert⁴ for instance, have equated biodiversity protection in the oceans to the establishment of no-take areas, writing: "Even lumping all categories together, only 3.5% of the ocean is protected" and "only 1.6% is 'strongly' or 'fully' protected." And in 2014, Carissa Klein and co-authors⁵ evaluated the degree to which the ranges of more than 17,000 species are contained within MPAs. I interpret this as implying that species whose ranges do not fall within MPAs will be lost, although these authors concede that, for some species, "the best conservation outcome may be achieved with other strategies, including fisheries regulations".

MANAGEMENT STRATEGIES

There are many other useful tools and legal frameworks designed to reduce overfishing, rebuild fish stocks and protect the biodiversity of the oceans. National and international fisheries agencies have been developing and enforcing these for the past two decades.

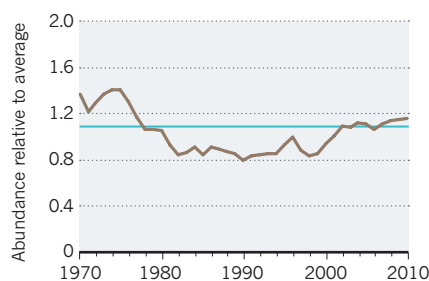
Problems are identified and tools selected to solve them in what is often a highly participatory process involving many stakeholders. If a certain fishing approach, such as bottom trawling, threatens a habitat, the area can be closed to that type of fishing. If

THE FRUITS OF FISHERIES MANAGEMENT

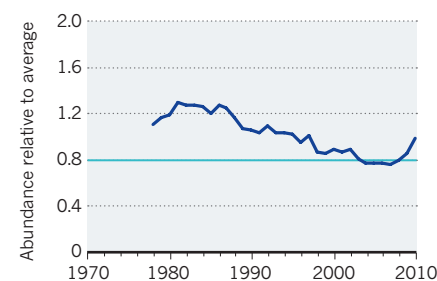
In several areas, fish stocks are stabilizing (A, C) or increasing (B, D) thanks to catches being limited or certain types of fishing gear being restricted in some areas.

— Amount of fish needed to produce long-term maximum sustainable yield

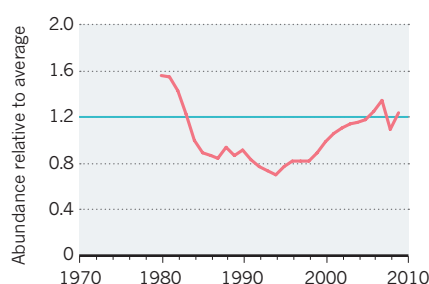
A Iceland and Norway



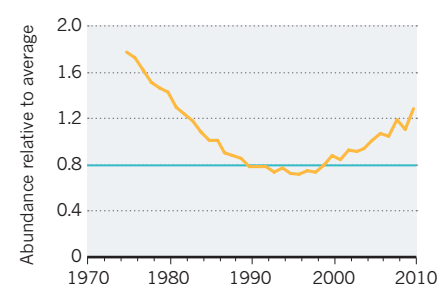
B EU Atlantic



C US East Coast



D US Southeast and Gulf of Mexico



a species is being threatened as a result of being caught unintentionally along with the targeted species, the fishery may be closed, fishing permitted at only certain times of the year, or catching techniques modified to reduce by-catch. Dolphin mortality fell almost 100-fold between 1986 and 1998 in the eastern Pacific⁶, for instance, after vessels changed fishing practice so that ensnared dolphins were released before the nets were hauled aboard. (The technology was developed by fishermen after the Inter-American Tropical Tuna Commission instituted limits to dolphin by-catch.)

The United States spends more than US\$300 million per year on fisheries management. It does so through the implementation of key pieces of legislation, including the Magnuson-Stevens Fishery Conservation and Management Act, the Endangered Species Act, the Marine Mammal Protection Act, the National Environmental Policy Act and the Clean Water Act. In Alaska, for example, more than 50% of the continental shelf waters are closed to specific kinds of fishing gear and the entire shelf is covered by species-specific catch restrictions. This is much more protection than could be offered by turning 30% of the region into MPAs.

"Biodiversity protection and fisheries management must be overseen by the same bodies."

Because of fisheries management, overfishing has largely been eliminated in US waters⁷. The proportion of fish stocks listed as 'overfished' — those in which abundance is lower than that needed to produce near-maximum yield — halved between 1997 and 2014 to 16% (see go.nature.com/2946lg4). Overfishing has also largely stopped in the European Union's Atlantic fisheries, New Zealand, Australia, Iceland, Norway and Canada (see 'The fruits of fisheries management')⁸. And management strategies recently implemented by major Latin American countries, including Peru, Argentina and Chile, have reduced the proportion of stocks that are fished above optimal rates from 75% in 2000 to 45% in 2011 (unpublished data).

In short, it is now clear that for those countries with effective fisheries management in place — a group of nations responsible for 45% of the global catch — fish stocks are stable, or increasing. Of course, most of the world's fisheries, especially in Africa and in parts of Asia, have no protection of any kind.

BRIDGE THE DIVIDE

Studies show that enforcing the closure of an area to fishing increases the density of fish in the reserve by around 166%¹. Yet, at best, MPAs will cover a small fraction of the ocean and few studies have evaluated their effect on biodiversity outside their perimeters. Catch



In the Philippines, communities are using a mix of protected areas and other forms of regulation to make fishing more sustainable.

data, records of boat movements and other monitoring efforts indicate that fishing pressure may increase beyond MPAs².

More pressingly, neither MPAs nor fisheries management alone can shield marine biodiversity from the panoply of current threats: climate change and ocean acidification, land-based run-off, oil spills, plastics, ship traffic, tidal and wind farms, ocean mining and underwater communications cables.

The enormity of the challenge calls for a change in approach. Instead of working at cross purposes, MPA advocates and those in fisheries management need to identify and solve area-specific problems together, and in consultation with diverse stakeholders. These may range from professional and recreational fishermen, park officers and environmental NGOs to developers, oil and gas companies and communications companies.

Regional coastal-management agencies, such as the California Coastal Commission, which operates as a quasi-independent government agency, are a potential model. But their mandate and membership would have to be significantly expanded if they were to deal with the impacts of fisheries and the establishment of MPAs. Such commissions have traditionally been confined to nearshore waters and have been able to regulate only development permits.

Marine spatial planning is a generic term for the process of resolving conflicts in the use of marine resources and would seem to be the obvious mechanism to integrate fisheries management and MPAs. Yet after

more than a decade of discussion and some attempts at implementation, there are few examples of the process effectively bringing the two 'tribes' together to work towards common goals. I suspect that this is, in part, because insufficient efforts have been made to convince both parties that decision-making bodies represent their interests appropriately.

The best examples of MPA advocates and fisheries-management communities working together are small-scale. In the Philippines and Indonesia, for instance, communities are working with local governments and NGOs, using a mix of protected areas and other forms of regulation, to try to rebuild coral-reef fish stocks⁹. Here the principal aim is to make fishing more sustainable; the objective of protecting representative habitats is not typically considered.

In larger industrial fisheries, such as in Europe, Australia and New Zealand, it should be possible for MPA advocates to collaborate with national fisheries departments. This would require a clear elaboration of the objectives of each. It would also require the appointment of more conservationists and MPA advocates to fisheries-management organizations, which are currently dominated by regulatory agencies and fishing-interest groups.

Another way to foster collaboration on a national scale would be to merge the various government departments responsible for conservation and fisheries management into a single department of marine management.

Such an organization could oversee the protection of biodiversity and the sustainable use of fisheries, and regulate competing marine uses. As a first step, a set of formal consultations, informed by case studies that measure the actual level of biodiversity protection achieved in different places through existing mixes of MPAs and fisheries management, could begin to identify clear measurable objectives.

At the local, national and international levels, biodiversity protection and fisheries management must be overseen by the same bodies if either is to be truly effective. ■

Ray Hilborn is professor in the School of Aquatic and Fisheries Sciences at the University of Washington, Seattle, USA. e-mail: rayh@uw.edu

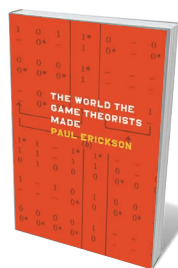
1. Lester, S. E. *et al.* *Mar. Ecol. Progr. Ser.* **384**, 33–46 (2009).
2. Hilborn, R. *et al.* *Ocean Coast. Mgmt* **47**, 197–205 (2004).
3. Lubchenco, J., Gaines, S. D., Warner, R. R., Airame, S. & Simler, B. *The Science of Marine Reserves* (Partnership for Interdisciplinary Studies of Coastal Oceans, 2002); available at <http://www.piscoweb.org>
4. Lubchenco, J. & Grorud-Colvert, K. *Science* **350**, 382–383 (2015).
5. Klein, C. J. *et al.* *Sci. Rep.* **5**, 17539 (2015).
6. Hall, M. A., Alverson, D. L. & Metuzals, K. I. *Mar. Pollut. Bull.* **41**, 204–219 (2000).
7. Ocean Studies Board. *Evaluating the Effectiveness of Fish Stock Rebuilding Plans in the United States* (National Academies Press, 2014).
8. Costello, C. *et al.* *Science* **338**, 517–520 (2012).
9. Christie, P., Campbell, L. & Armada, N. in *Governance of Marine Fisheries and Biodiversity Conservation* (eds Garcia, S. M., Rice, J. & Charles, A.) 332–345 (Wiley-Blackwell, 2014).



ILLUSTRATION BY TOM CLOHOSY COLE

SUMMER BOOKS

As labs and lecture halls empty, go out of this world with our regular reviewers' recommendations for stellar holiday reading.



The World the Game Theorists Made

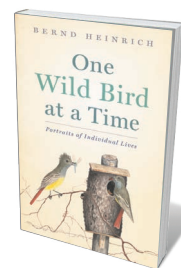
PAUL ERICKSON
Univ. Chicago Press: 2015.

In this deeply researched and readable book, historian Paul Erickson chronicles the passage of game theory from mathematical economics to arms-control theory to evolutionary biology, and back to economics. Along the way, it has changed considerably, and now is dominated almost entirely by “non-cooperative” games that obey the equilibrium laid out by mathematician John Nash.

Until now, there has been remarkably little written about the history of game theory since its creation in 1944 by mathematician John von Neumann and economist Oskar Morgenstern, beyond recounting a storied intellectual lineage studded with economics Nobel laureates. Such accounts do little to explain how game theory came to be applied to the novels of Jane Austen (among much else). Some concepts — the prisoner’s dilemma, the tragedy of the commons — are today almost impossible to discuss without someone scribbling down a 2×2 grid.

Erickson — with whom I collaborated on *How Reason Almost Lost Its Mind* (Univ. Chicago Press, 2013), a history of cold-war rationality — brings a distinctive voice to the material. The book will edify and surprise even specialists, let alone the rest of us, who have to live in a world created by game theorists whether or not we can articulate the pay-offs in the matrix.

Michael D. Gordin is a historian at Princeton University in New Jersey.



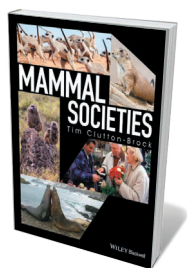
One Wild Bird at a Time: Portraits of Individual Lives

BERND HEINRICH
Houghton Mifflin Harcourt: 2016.

Here, renowned biologist Bernd Heinrich shows us how to find adventure in the nature in our neighbourhood. Heinrich examines the behaviour of the wonderful birds — from flycatchers to owls — that throng the vicinity of his Maine cabin. He invites us to witness his well-honed scientific process as he demystifies seemingly inexplicable behaviours, from crow-on-crow murder, sapsucker drumming and the ‘snow caves’ of grouse to some raptors’ predilection for adding greenery to their nests.

The book inspires us to become backyard sleuths. Some solutions require more than patience and a keen naturalist’s eye. The northern flicker (*Colaptes auratus*), for example, nests in cavities and so is rarely visible — although not to Heinrich, who chainsaws a hole in his cabin wall to install an observable nesting box. This home ‘improvement’ project, along with Heinrich’s insights into the ethological riches to be found in forest and field, will propel you outside this summer, to follow the denizens of your own patch of the wild.

John M. Marzluff is the James W. Ridgeway Professor of Wildlife Science in the School of Environmental and Forest Sciences at the University of Washington, Seattle. His latest book is *Welcome to Subirdia*.



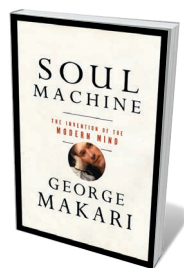
Mammal Societies

TIM CLUTTON-BROCK
Wiley-Blackwell: 2016.

Mammals come in all shapes and sizes, but they share key features of their reproductive biology. For instance, females give birth to live young and produce milk to nourish them. The costs of gestation and lactation shape the females' reproductive strategies, and this, in turn, influences the reproductive strategies of the males.

Over the past 40 years, behavioural ecologists have documented how these dynamics play out in a range of species. Tim Clutton-Brock has had a leading role in this endeavour, conducting long-term studies of Scottish fauna (red deer — *Cervus elaphus* — on the island of Rum, and Soay sheep in the St Kilda archipelago) and meerkats (*Suricata suricatta*) in Africa's Kalahari Desert. In *Mammal Societies*, he provides a masterful synthesis of what we know about the behavioural ecology of mammals. Clutton-Brock skilfully weaves an immense body of material into a coherent narrative that emphasizes the processes that shape adaptation, and reveals common patterns that have evolved. The text is gracefully written, enlivened by well-chosen examples and beautifully illustrated. (A few of my own photographs of baboons appear in it.) This is a book that Charles Darwin would have wanted in his library; it should be in yours.

Joan B. Silk is at the Institute of Human Origins at Arizona State University, Tempe.



Soul Machine: The Invention of the Modern Mind

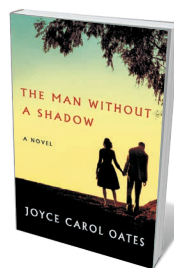
GEORGE MAKARI
W. W. Norton: 2015.

In this brilliant mixture of history, philosophy and science, psychiatrist and historian George Makari explores the origins of our ideas about self and that ephemeral phenomenon, the mind. Focusing on the 'long eighteenth century' (1660–1830), Makari explores how early-modern thinkers and Enlightenment philosophers shaped the way we think about reason, knowledge, language and consciousness.

Makari traces disputes between philosophers René Descartes, Thomas Hobbes and John Locke in this arena, vividly linking modern neuroscience with thinkers whom many scientists may not have read. He also looks at changing attitudes to mental health, describing how the York Lunatic Asylum in England chained up naked inmates, leading to the death of a Quaker patient in 1790. In response, Quakers set up the York Retreat, with its moral and behavioural treatments — a model that had gained broader support by 1814.

Insightful, thought-provoking and encyclopaedic, Makari's book shows how we came to understand where the mind is located and something of its nature. He closes on science's present impasse: the inability to properly explain consciousness or what happens when mental illness affects us. We remain, he argues, "modern hybrids of soul and machine".

Matthew Cobb is professor of zoology at the University of Manchester, UK.



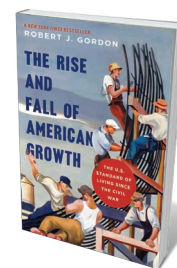
The Man Without a Shadow: A Novel

JOYCE CAROL OATES
Ecco: 2016.

The acclaimed US writer Joyce Carol Oates's 44th novel follows the career of fictional neuropsychologist Margot Sharpe and her interactions with "E.H.", a man with anterograde amnesia who can hold on to newly made memories for a mere 70 seconds. Charismatic and clever, but perennially confused, E.H. can recall only events that happened before his brain was damaged, and is haunted by vivid, mysterious recollections of a drowned girl. Sharpe is far from normal herself: friendless and work-obsessed, she struggles to succeed as one of the few female scientists in her field.

Reflecting real-life research, the story borrows from famous experiments used by neuroscientists such as Suzanne Corkin in her studies of celebrated amnesiac patient Henry Molaison. The claustrophobic action takes place over decades, almost entirely in Sharpe's research lab. As Sharpe is slowly drawn over the ethical line in her obsession with E.H., we witness her triumph and shame. In illuminating the science of memory, this uneasy thriller makes us question how much our capacity to revisit the past allows us to contemplate the future.

Jennifer Rohn leads a cell-biology group at University College London. Her most recent novel is *The Honest Look*.



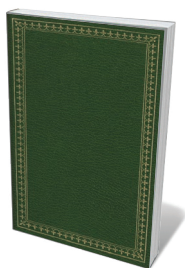
The Rise and Fall of American Growth: The US Standard of Living Since the Civil War

ROBERT GORDON
Princeton Univ. Press: 2016.

The must-read book of the moment for economists is Robert Gordon's *The Rise and Fall of American Growth*. The adjective 'magisterial' applies (making it perhaps too hefty for beach reading). The first half is a fine, detailed description of the dramatic increase in US living standards, which mirrors those of other industrializing economies, from the mid-nineteenth to late twentieth centuries. Gordon concentrates on the role of technological progress, particularly innovations in public health and transportation, as the main driver of growth. In the second half, Gordon airs his more controversial belief that today's advance in digital technologies is much less important, and amounts to a technological slowdown that has flat-lined productivity. However, he has narrowly (and thus unpersuasively) focused on too few technologies.

More convincing on the gradual adoption of innovations in industry, and their influence on jobs and living standards — again with plenty of history — is James Bessen's excellent and comparatively optimistic *Learning By Doing* (Yale Univ. Press, 2015).

Diane Coyle is a professor of economics at the University of Manchester, UK, and author of *GDP: A Brief but Affectionate History*.



The Descent of Man

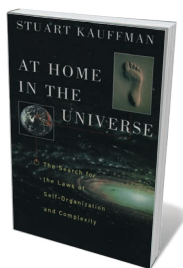
CHARLES DARWIN
John Murray: 1871.

In my archaeological career, I often find myself taken right back to Charles Darwin. A recent visit to the Museum of Mankind in Paris was one such moment. Touring the human-evolution displays in the new permanent galleries, which reopened last year, I found that aside from twenty-first-century twists such as an explanation of the human genome, the exhibits are in many respects like an illustrated edition of Darwin's 1871 *The Descent of Man*.

In this book, Darwin first establishes that humans carry indelible signs of our origins, such as physiological similarities to the great apes. He then sets out to show that, despite representing a minuscule part of the living world, we have had the greatest impact. Crucially for archaeologists, he identifies the elements that make us human as culture, mind and conscience. Finally, he puts us right back in our biological place by attributing these distinctions to sexual selection.

Arguments about selection are now more nuanced, but who we are, where we come from and where we are going remain fascinating questions for philosophers and scientists. Darwin's Victorian language about women and race is outmoded, but his ideas set the course for debates on what defines humanity, while incorporating instructive observations from nature, from tool use to nest building.

Jill Cook is acting keeper in the Department of Britain, Europe and Prehistory at the British Museum in London.



At Home in the Universe: The Search for the Laws of Self-Organization and Complexity

STUART KAUFFMAN
Oxford University Press: 1995.

If we could rewind the tape of life to its inception around 3.6 billion years ago, and press the 'play' button, would humans emerge once more? Are we the expected and inexorable products of evolutionary processes, or are we chance occurrences that might never evolve again? If natural selection were the only organizing principle that determined the structures of living things, it is unlikely that the historical pageant of evolutionary form would recur.

More than two decades ago, in *At Home in the Universe*, theoretical biologist Stuart Kauffman argued that life's history has been guided and constrained by an additional source of order: the self-organizing and emergent properties of complex, far-from-equilibrium systems poised at the edge of chaos. The well of natural order that issues from such critically tuned complex systems may be driven by a set of deep laws that invisibly guide the evolution of life on Earth, and presumably throughout the Universe. Kauffman explores these by modelling the behaviour of highly interconnected Boolean networks. His prescient observations may have a new relevance in the era of genome editing and microbiome dynamics.

Adrian Woolfson is the author of *Life Without Genes*.



ILLUSTRATION BY TOM GLOHOSY COLE

Correspondence

Israel must release Palestinian physicist

On 24 April, the distinguished Palestinian astrophysicist Imad Ahmad Barghouti was arrested and detained without charge by the Israeli military — for the second time in less than 18 months (see *Nature* <http://doi.org/bk44>; 2016). We protest against his imprisonment and renew the call for his release.

Once again, it was alleged that Barghouti made statements on Facebook and on television in opposition to Israeli military attacks and occupation. Subsequent international pressure again contributed to a ruling by the military court of appeals a month later for his release (see go.nature.com/299v9nd). This has not happened. Barghouti has instead been transferred to a facility run by Israel's internal security service, Shin Bet, for further interrogation.

Ahmed Abbas* *IHES, CNRS, Paris, France.*
abbes@ihes.fr

*Supported by 13 signatories (see go.nature.com/29btthb).

Build social costs into wildfire risk

Marc-André Parisien calls for the fire-science community to draw up maps of wildfire risk based on contributory factors such as local topography, vegetation and weather (*Nature* **534**, 297; 2016). I suggest that incorporating social and economic risks into such maps would increase the effectiveness of fire-management policies.

The insurance and housing industries are likely to discourage construction of fire-sensitive structures in zones designated as high risk in wildfire maps. We therefore need proper accounting of the implications of fire-risk mapping for zoning, taxation and insurance fees. For example, introducing these maps could affect rural communities, which may have to pay higher premiums or be denied fire insurance.

We do not yet fully understand how to project human-caused fire risk. Demographics may explain global shifts in pyrogeography (W. Knorr *et al.* *Nature Clim. Change* <http://doi.org/bkqn>; 2016); changing land use and agricultural practices also contribute.

National fire-risk mapping systems will need to take into account local differences in governance and attitudes if they are to be effective at regional or community levels. There may be mismatches in ecosystem vulnerability, fire policies and acceptable degrees of fire risk. In the United States, for example, prescribed burning for habitat restoration is relatively new in Pennsylvania, but long established in neighbouring New Jersey.

Erica A. H. Smithwick
Pennsylvania State University, University Park, USA.
smithwick@psu.edu

Entice Africa's scientists to stay

The Dakar Declaration was put forward at the 2016 global meeting of the Next Einstein Forum held in Dakar, Senegal. It focuses on advancing African science, technology, engineering and mathematics (STEM), with particular encouragement for women and young scientists (go.nature.com/29ra0k6; see also *Nature* **531**, 275; 2016). Previous attempts to create a scientific identity for Africa have met with limited success, so it is crucial to improve the continent's research environment if we are to realize and maintain its potential.

In many African countries, this research environment is characterized by sparse financial, human and capital resources, lack of opportunities and inadequate infrastructure. Such conditions drive many young and talented scientists to leave for other parts of the world.

In our view, plans to move STEM forward in African

countries should focus on strategic utilization of scientific research for economic and societal development. Restructuring of the governance of science and technology is essential for a knowledge-based economy. The call to invest more in promising young African scientists should encourage their reintegration into their home countries after working and studying abroad — and prevent a brain drain in the long term.

Ghada Bassioni *Ain Shams University, Cairo, Egypt.*
Gameli Adzaho *Weija, Ghana.*
David Niyukuri *University of Burundi, Bujumbura, Burundi.*
ghada_bassioni@eng.asu.edu.eg

Speed restoration of EU ecosystems

The European Union's Biodiversity Strategy 2020 aims to restore at least 15% of degraded ecosystems by 2020, in accordance with the 2010 Aichi targets. With these due for review later this year, we are deeply concerned about Europe's lack of progress towards meeting its own target. In our view, EU member states need to engage properly in the repair of their damaged ecosystems.

All member states failed to honour their commitment to deliver a sound national restoration prioritization framework by the end of 2014, an initiative intended to improve the quality, scale and consistency of ecosystem restoration. This was profoundly disappointing, given the poor state of the environment in large parts of Europe and the need to develop resilience against climate change.

We urge the European Commission and member states to rectify this potentially costly inaction. Ecosystem restoration must be fully integrated into EU and national policies and land-use planning. Fact-based guidelines and regulations should be issued and standards set up for evaluation and progress. Provision of additional

earmarked funds is essential to boost cost-effective restoration.

Countries with a robust restoration plan and confirmed restoration efforts could be rewarded with access to further EU funding. It is also important to publicize the benefits of restoration for social well-being.

Jordi Cortina-Segarra *University of Alicante, Spain.*
Kris Decler *Research Institute for Nature and Forest, Belgium.*
Johannes Kollmann *Technical University of Munich, Germany.*
jordi@ua.es

Citizens arrest river pollution in China

We would like to share the story of how local citizens have helped to clean up polluted rivers in Zhejiang province in China.

Three years ago, the regional government allocated 140 billion yuan (US\$21.5 billion) to the clean-up operation. It launched a campaign to alert the local population of 55 million to the severity of the pollution and the risk to public health, and to publicize its plans and the benefits in the long term. As a result, Zhejiang citizens support the project and voluntarily cooperate to implement it.

Individual rivers are overseen by a government official and a technology expert, whose contact details are posted at the riverside so that the public can promptly report illegal polluted discharges from factories and other sources. The officers then order these discharges to be cut off, knowing that their own promotion depends on swift action and improvements in water quality.

This overall strategy has markedly increased urban water quality (go.nature.com/292hwbi; in Chinese). It has also enabled the government to economize on monitoring equipment.
Tuqiao Zhang, Feifei Zheng, Tingchao Yu *Zhejiang University, Hangzhou, China.*
feifeizheng@zju.edu.cn

Jerome Bruner

(1915–2016)

Psychologist who shaped ideas about perception, cognition and education.

Jerome Seymour Bruner helped to launch the cognitive revolution in psychology — the shift from focusing on how stimuli or rewards provoke behaviours (behaviourism) to trying to understand the workings of the mind.

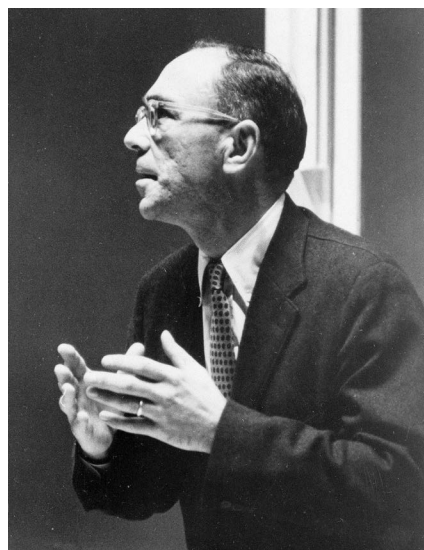
Bruner, who died on 5 June, aged 100, was born — blind — in New York City in 1915. His sight was restored by cataract operations when he was two. In 1937, he earned a degree in psychology at Duke University in Durham, North Carolina. He received master's and doctoral degrees in psychology from Harvard University in Cambridge, Massachusetts, in 1939 and 1941. After working in military intelligence during the Second World War, he took a faculty position at Harvard in 1945.

Bruner once noted that during his two years of blindness, he had constructed a visual world in his mind. His early experiences may explain why, in the 1940s and 1950s, he sought to demonstrate how perception is not just a bottom-up process controlled by the senses, but also a top-down process controlled by the mind.

In collaborative experiments conducted at Harvard, Bruner revealed how certain mental factors influence visual perception. In one study, for instance, he demonstrated that ten-year-old children overestimate the size of bigger coins and underestimate the size of smaller coins, and that poor children overestimate the size of the larger coins more than affluent ones do (J. S. Bruner *et al.* *J. Abnorm. Soc. Psychol.* **42**, 33–44; 1947). His work inspired a new approach to the study of perception that became known as the new look in perception.

Bruner transformed perception from a stimulus-dependent response into something that involved mental processing. But he wanted to study cognition more directly. With psychologists Jacqueline Goodnow and George Austin, he performed innovative experiments that explored how people infer concepts and categories (for instance, of colour and shape). Their 1956 book *A Study of Thinking* was crucial in ushering in the cognitive revolution.

Bruner's 1960 book, *The Process of Education* brought the cognitive revolution to educational thinking in the United States and elsewhere. His concepts of the development of representational capacities, suggested that ideas should be communicated to students using actions, icons or symbols, in that order, and depending on their age. In 1963, after



my first year in graduate school at Harvard, Bruner arranged for me to go to Senegal to study culture and cognitive development. Because his book had tightly linked schooling with cognitive development, Bruner was delighted when my data from Senegal showed that various measures of such development depended on whether or not children had attended school, not just their age.

I had first encountered Bruner during my first year of university, when he lectured in one of my courses. In his lectures, he described the concept of human intentionality — the ability of the mind to be proactive and to represent future goals — as another challenge to behaviourism. When I returned to Harvard as a research fellow in 1968, Bruner was studying cognitive development in infancy. Intentionality was now central to his thinking. Crucially, he observed that infants only a few weeks or months old have intentions and goals, even before they are able to act on them.

In 1972, Bruner sailed his boat across the Atlantic to take up the first Watts Professorship of Psychology at the University of Oxford, UK. There, he shifted focus from 'intentional action' to 'intentional interaction'. In 1975, Michael Scaife and Bruner reported that starting at eight months old, most infants will follow an adult's gaze when the adult turns to look at something (M. Scaife and J. S. Bruner *Nature* **253**, 265–266; 1975). The pair called this phenomenon joint visual attention because it established a common focus between adult and infant. It has since

become widely recognized as an essential social mechanism for guiding infants to link words to objects in language acquisition.

Bruner returned to the United States in 1980. First at the New School for Social Research in New York, then at New York University, he explored people's propensity to tell stories. He argued that unlike logic, narrative thought is universal. Once again, he was trying to expand cognitive psychology to encompass human experience. His 1986 book on narrative, *Actual Minds, Possible Worlds*, has more than 14,100 citations. Applying ideas about narrative to the law, Bruner started working with legal scholar Anthony Amsterdam at the New York University School of Law. Their 2000 book *Minding the Law* describes how courts rely on storytelling and how the stories "change the way we understand the law — and ourselves".

Jerry made seminal contributions to an astonishing number of fields — each a stop on the road to finding out what makes us human. Beginning in the 1960s, computer simulations became the model of the human mind in cognitive psychology, with researchers trying to simulate how humans solve problems, form concepts, comprehend language and learn. But reducing humans to computers was antithetical to Jerry's humanistic perspective.

Given this, it was surprising that computer scientist Alan Kay, the designer of what became the Macintosh graphical user interface, turned up more than 30 years ago on Bruner's Manhattan doorstep with a gift of a Macintosh computer. Jerry's ideas of representing information through actions, icons and symbols, central to his theory of cognitive development, had inspired Kay to get users (even children) to act (through a computer mouse) on icons, enabling the use of an abstract set of symbols (computer program). This was the foundation for what became the Macintosh interface.

Jerry had a towering intellect and an insatiable curiosity. When I returned from Senegal with my data, he made me feel as if I had done the most exciting research in the world. His reaction fuelled the rest of my career and has greatly influenced my own mentoring. ■

Patricia Marks Greenfield is distinguished professor of psychology at the University of California, Los Angeles, USA. Jerome Bruner was her teacher, mentor, colleague and lifelong friend.
e-mail: greenfield@psych.ucla.edu

natureOUTLOOK

14 July 2016
Supplement to Nature
Publishing Group journals

PAIN



Produced with support from:

ESTEVE



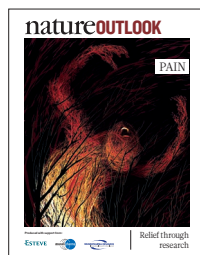
MUNDIPHARMA RESEARCH
International Research & Development

Relief through
research

natureOUTLOOK

PAIN

14 July 2016 / Vol 535 / Issue No 7611



Cover art: Daniel Hertzberg

Editorial

Herb Brody
Michelle Grayson
Richard Hodson
Jenny Rooke

Art & Design

Mohamed Ashour
Andrea Duffy
Wesley Fernandes

Production

Matthew Carey
Ian Pope
Karl Smart

Sponsorship

Stephen Brown
Samantha Morley

Marketing

Nicole Jackson

Project Manager

Anastasia Panoutsou

Art Director

Kelly Buckheit Krause

Publisher

Richard Hughes

Editorial Director, Partnership Media

Stephen Pincock

Chief Magazine Editor

Rosie Mestel

Editor-in-Chief

Philip Campbell

Pain is unpleasant but necessary. It signals danger, preventing us from harming ourselves, and alerts us to damage to our bodies. Yet for many people, their pain system is out of alignment. Too much pain is crippling and can make everyday living an agony. Even 'good' pain can turn bad, when the pain of an injury persists after the damage has healed.

To use only the term 'pain' is to ignore the full range of ways that people experience hurt. It might all start with the same basic pathways (see page S2), but the nuances change as acute pain becomes chronic and even the strongest analgesics stop working (S4). Damage to nerves has its own set of effects. Which sensation a person experiences might be a clue to the cause of their neuropathic pain — and how to treat it (S10).

The personal nature of pain complicates its study. Men and women even process pain through different immune cells in the spinal cord. Such an important distinction has a bearing on the sex of animals used in pain research (S7). And these variations have hampered genetic studies, which have so far shown only that pain is mediated by a mosaic of thousands of genes (S12). Brain imaging, however, is providing more leads. Researchers think that they have identified a neurological signature of pain that could be used in comparison studies (S8).

Researchers have come a long way in terms of understanding and controlling pain (S18). But although people in developed countries have access to the strongest opioids, billions elsewhere do not — even those in palliative care (S16). However, things are changing, with many drugs and devices in development, including the intriguing possibility of using honest placebos as painkillers (S14).

We are pleased to acknowledge the financial support of ESTEVE, Mundipharma International Ltd and Mundipharma Research GmbH & Co. KG in producing this Outlook. As always, *Nature* has sole responsibility for all editorial content.

Michelle Grayson

Senior editor, supplements

CONTENTS

S2 NEUROSCIENCE

The pain drain

How good pain turns bad

S4 BIOMEDICINE

Move over, morphine

Alternatives may be around the corner

S7 PERSPECTIVE

Equality need not be painful

Jeffrey S. Mogil on using female subjects

S8 IMAGING

Show me where it hurts

Technology is revealing a pain pattern

S10 NEUROPATHY

A name for their pain

Categorizing pain to improve treatment

S12 GENETICS

An incomplete mosaic

Pinpointing a cause for chronic pain

S14 PLACEBOS

Honest fakery

The real response to fake pills

S16 PALLIATIVE CARE

The other opioid issue

Overcoming a fear of morphine use

S18 HISTORY

Painful progress

Research through the ages

RELATED ARTICLES

S20 Corticostriatal functional connectivity predicts transition to chronic back pain
M. N. Baliki et al.

S23 Different immune cells mediate mechanical pain hypersensitivity in male and female mice
R. E. Sorge et al.

S26 Cross-modal representations of first-hand and vicarious pain, disgust and fairness in insular and cingulate cortex
C. Corradi-Dell'Acqua et al.

S38 NIH to balance sex in cell and animal studies
J. A. Clayton & F. S. Collins

S40 A thalamic input to the nucleus accumbens mediates opiate dependence
Y. Zhu et al.

Nature Outlooks are sponsored supplements that aim to stimulate interest and debate around a subject of interest to the sponsor, while satisfying the editorial values of *Nature* and our readers' expectations. The boundaries of sponsor involvement are clearly delineated in the *Nature Outlook* Editorial guidelines available at go.nature.com/e4dwzw

CITING THE OUTLOOK

Cite as a supplement to *Nature*, for example, *Nature* Vol. XXX, No. XXXX Suppl., Sxx–Sxx (2016).

VISIT THE OUTLOOK ONLINE

The *Nature Outlook Pain* supplement can be found at <http://www.nature.com/nature/outlook/pain>. It features all newly commissioned content as well as a selection of relevant previously published material.

All featured articles will be freely available for 6 months.

SUBSCRIPTIONS AND CUSTOMER SERVICES

Site licences (www.nature.com/libraries/site_licences): Americas, institutions@natureny.com; Asia-Pacific, <http://nature.asia/jp-contact>; Australia/New Zealand, nature@macmillan.com.au; Europe/ROW, institutions@nature.com; India, ngindia@nature.com. Personal subscriptions: UK/Europe/ROW, subscriptions@nature.com; USA/Canada/Latin America, subscriptions@us.nature.com; Japan, <http://nature.asia/jp-contact>; China, <http://nature.asia/china-subscribe>; Korea, www.natureasia.com/ko-kr/subscribe.

CUSTOMER SERVICES

Feedback@nature.com

© 2016 Macmillan Publishers Limited. All rights reserved.

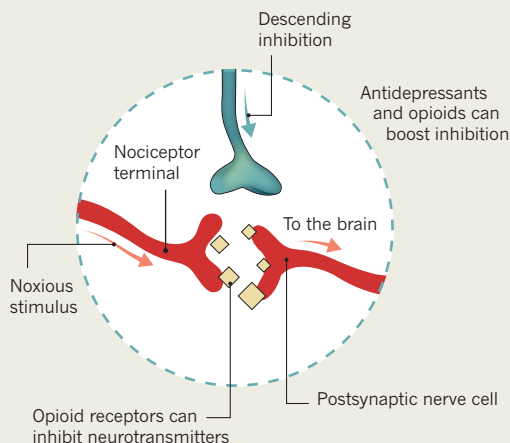
THE PAIN DRAIN

We can't live without it, but many of us struggle to live with it. Pain has an essential biological function, but too much — or the wrong sort — ruins lives and puts a sizeable dent in economic productivity. By **David Holmes**, infographic by **Mohamed Ashour**.

3

FEELING THE PAIN

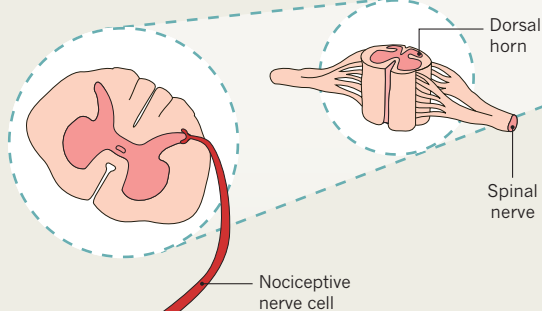
Nociceptive inputs to the dorsal horn are subject to powerful descending control from a network of areas in the brain. Signals combine in the brainstem, then travel down (blue arrows) to alter ascending nociceptive signals. Many painkillers, such as opioids and antidepressants, act at this junction.



2

SIGNAL

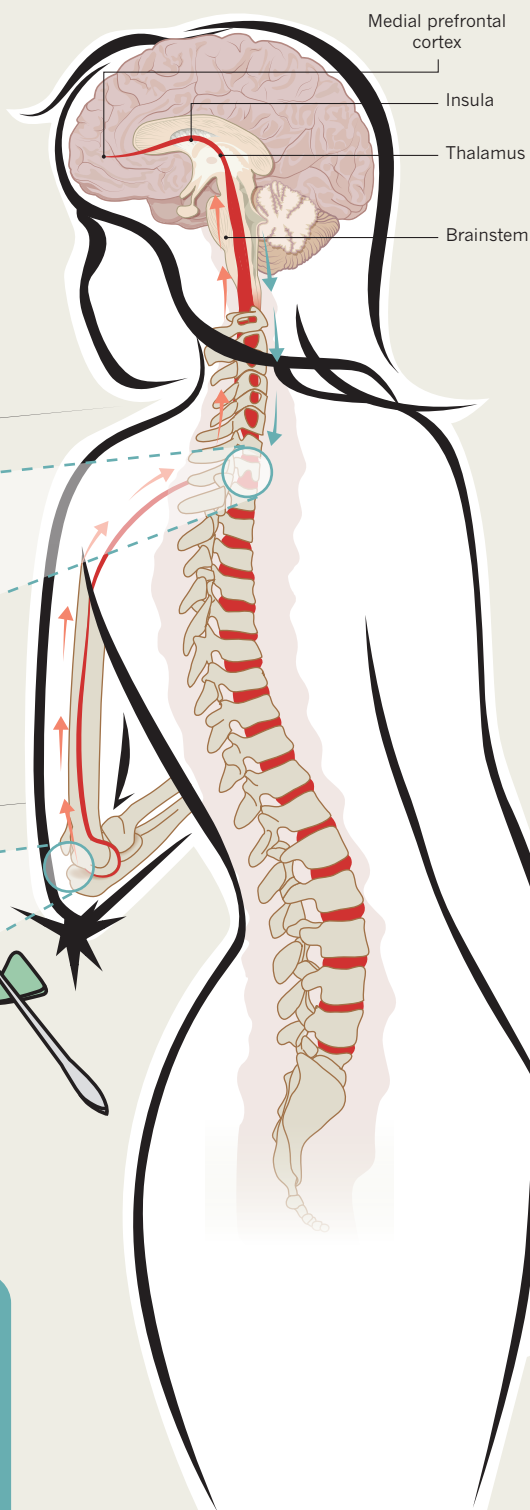
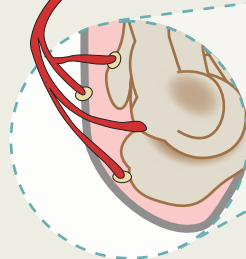
The pain signal is transmitted along axons of nociceptive nerve cells to the dorsal horn of the spinal cord. Signals from the dorsal horn then pass to the brain (red arrows).



1

TRAUMA

Nociceptive pain starts with the stimulation of nociceptors, which are found in the skin, internal organs, muscles, joints, and the membranes around the brain and spinal cord.



NOCICEPTIVE PAIN

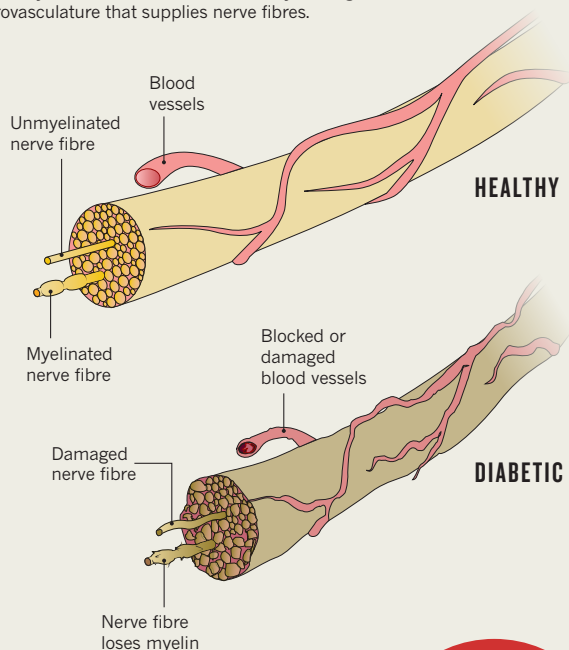
This type of pain is caused by the activation of nociceptors — specialized sensory neurons that are stimulated by noxious mechanical, thermal or chemical stimuli. Nociceptors transform these stimuli into electrical signals and relay them to the central nervous system. Nociceptive pain tends to be short-lived and associated with injury. But if it persists beyond 12 weeks, it becomes chronic pain — and its nature can change.

NEUROPATHIC PAIN

Unlike nociceptive pain, neuropathic pain is caused by damage to the somatosensory nervous system itself, as a result of trauma or disease. However, there is not always a clear link between disease states and neuropathic pain.

DIABETIC NEUROPATHY

Painful diabetic peripheral neuropathy is one of the most common forms of neuropathic pain, with its incidence set to increase as the obesity and diabetes epidemics continue to grow. Neuropathy is caused by metabolic factors as well as by damage to the microvasculature that supplies nerve fibres.

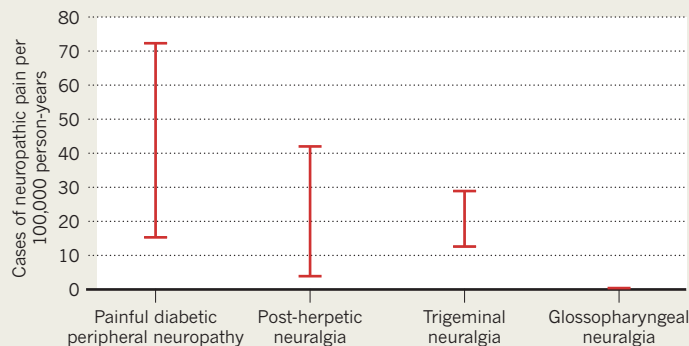


7–10%

Population prevalence of neuropathic pain¹

NEUROPATHIC PAIN INCIDENCE

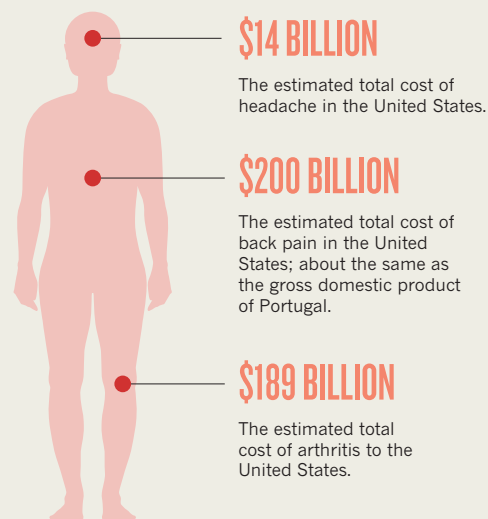
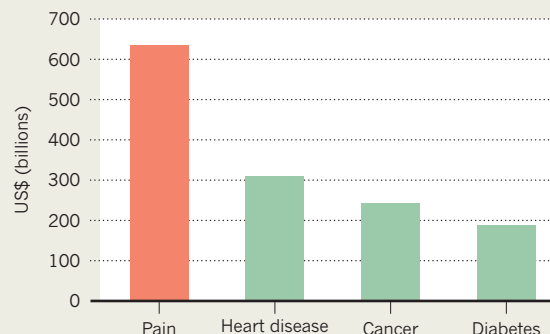
Definitions of neuropathic pain vary across studies, leading researchers to call for a unified nomenclature. The best evidence on incidence comes from studies of neuropathic pain linked to specific conditions, but even then ranges can vary widely¹.



PRICE OF PAIN

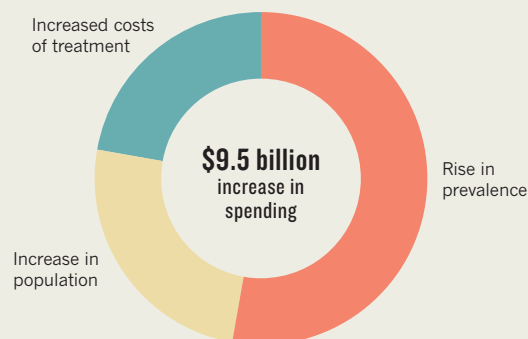
BIGGEST BURDEN

Around 100 million adults in the United States are affected by chronic pain in a single year. The annual total cost of pain, including direct costs, decreased wages and lost productivity, eclipses that of any other condition².



GROWING PAIN

Health-care spending on back problems in the United States more than doubled between 1987 and 2000. Although treatment costs and population increases contributed, most of the \$9.5-billion rise was due to an increase in the prevalence of back pain³.



1. van Hecke, O. et al. *Pain* **155**, 654–662 (2014). 2. Inst. Medicine *Relieving Pain in America* (National Academy of Sciences, 2011). 3. Thorpe, K. E. et al. *Health Affairs* <http://dx.doi.org/10.1377/hlthaff.w4.437> (2004).



Michael was 15 when he was kicked in the back by another student at his school in Australia. The blow ruptured a disc in his lower spine, a painful injury that required surgery. In the weeks and months that followed, Michael had additional operations, but none were able to resolve his excruciating pain.

After 12 rounds of surgery — at the end of which, three-quarters of Michael's back had been fused with rods and screws — the pain was undimmed. Michael (not his real name) had spent most of his adolescence in hospital and had become morbidly obese.

"When I first met him at 20, he was essentially bed-bound, virtually no sleep from the pain, on crazy doses of strong opioids," says Marc Russo, who directs the Hunter Pain Clinic in Newcastle, Australia. "He was existing, but certainly not living."

Unfortunately, cases such as Michael's are not rare. Globally, around one in five adults — almost a billion people — has persistent pain, which is often accompanied by sleep loss, depression, unemployment and relationship breakdown.

And for most people, the pain does not start with a trauma, as it did for Michael, but rather with a small physical insult, says Lorimer Moseley, a chronic-pain researcher at the University of South Australia in Adelaide. "Maybe they bent over to pick something up and their back hurts." A bout of pain — whether it's back strain or post-surgical pain — is often short-lived. But for around 10% of these people, the pain does not go away; they have transitioned from acute to chronic pain.

As Michael found out, effective treatments are sorely lacking. "There are a range of options, and a lot of people don't respond to any of them," says Charles Brooker, a pain-management specialist at the Royal North Shore Hospital in Sydney, Australia. So acute was the shortage of effective drugs, that some doctors — particularly in the United States, but also in other Western countries, including Australia — began prescribing strong opioids for chronic pain.

That strategy has turned out to be tragically misguided, leading to an epidemic of opioid addiction. "Opioids almost never work in chronic pain, and cause untold misery," says pain specialist Andrew Moore at the University of Oxford, UK. The good news is that as researchers finally begin to understand the mechanisms of chronic pain, new therapies — both drugs and devices — promise a powerful set of alternatives to opioids.

FROM KILLING TO CAUSING PAIN

Opioids kill pain by targeting μ -opioid receptors on pain neurons in the spinal cord and brain. By binding to these receptors, opioids inhibit neurotransmitter release at the junction between pain neurons, blocking

BIOMEDICINE

Move over, morphine

The dearth of treatment options for chronic pain has led to widespread over-prescription of strong opioids. But some innovative thinking is building a promising pipeline.

BY JAMES MITCHELL CROW

the signal. “Strong opioids are wonderful for palliative care and acute pain,” says Russo.

But for the long-term treatment of chronic pain, the side effects take a toll. The body quickly develops a tolerance, which doctors counter by escalating the dose. When Russo first began to see Michael for his back pain, one of his first interventions was to ease Michael off the opioids.

Understanding the mechanism of opioid tolerance may help researchers to find a way to avoid it. As more of the drug enters the body, non-neuronal cells known as glia take notice. Once thought to be nothing more than a scaffold for neurons, glia are now known to be active members of the central nervous system. One of the jobs of the glia is to keep watch for foreign invaders. High opioid doses seem to trigger a defensive response, causing the glia to release immune-signalling compounds called inflammatory cytokines, which stimulate the sensory neurons that the drugs are supposed to sooth. “The dose is so large it is now causing pain,” Russo says.

But because μ -opioid receptors are such powerful targets for suppressing pain, research into new opioids has not been abandoned entirely. Researchers are looking for drugs that weakly activate the μ receptor, but also hit other targets, says Russo. Hitting the μ receptor disrupts the flow of pain signals to the brain. Ideally, dual-acting drugs would also activate inhibitory nerves that descend from the brain to suppress pain, he says. This is how two of the newer morphine-derived drugs, tramadol (approved in the United States in 1995) and tapentadol (approved in 2008), work.

James Zadina, a neuroscientist who studies novel opioids at Tulane University in New Orleans, Louisiana, has taken a different approach. “Instead of starting from the opium plant, we started from the brain,” he says. The first big break came in 1997, when his team finally tracked down a peptide in the brain that binds to the μ receptor just as selectively as morphine does¹. The compound, called endomorphin-1, is now recognized as the natural trigger of the μ receptor.

Endomorphin-1 elicited an analgesic response just as powerful as morphine, but without the side effects. “The old way of thinking was, any drug that hits that receptor is going to do pretty much the same thing,” Zadina explains. In reality, however, drug molecules of different shapes can bind to the same receptor in different ways and trigger a different set of responses — a phenomenon known as biased agonism. In the case of the μ receptor, endomorphin-1 seems to selectively promote analgesia.

As drugs, natural endomorphins would be hopeless — they break down too rapidly in the bloodstream. Zadina and his

colleagues have been testing endomorphin analogues with reinforced chemical structures. The hope is that these molecules will still trigger the same response as the parent compound. Zadina has identified four new analogues of endomorphin², and is now

“There are a range of options, and a lot of people don’t respond to any of them.”

preparing to take the best-performing into clinical trials. That compound, dubbed analogue 4, provides “much longer duration of analgesia” than morphine, says Zadina. Tolerance is also reduced — and the compound does not seem to trigger the release of pain-stimulating cytokines.

In addition, analogue 4 does not seem to be addictive. The most compelling data, Zadina says, come from trials in which a rat can press a bar to self-administer the drug. A rat given access to morphine, he says, “will start pushing the bar like crazy. They don’t do that for our compound.”

CHRONIC-PAIN PROPHYLAXIS

However effective these pharmacological interventions prove to be, prevention will always be preferable. All chronic pain starts as acute pain. “It would be far better to extinguish it at origin,” says Russo.

Several studies have pinpointed factors that predispose patients to chronic pain — susceptibilities that a simple questionnaire can flag. Pre-existing anxiety and depression put people at risk, as does pre-existing pain. “People with chronic migraine are more likely to get chronic knee pain after a knee operation,” says Brooker. “Those people have a sensitized nervous system.”

Multiple sensitization mechanisms could be at work. There is a simple test to see whether one particular gate in the pain pathway, called diffuse noxious inhibitory control (DNIC), is functioning. For most people, if you place their left hand in ice water, they don’t feel mild pain induced by a laser shone on their right hand — the DNIC ‘filter’ in their spinal column is helping the brain to prioritize its response so that the person pays attention to the stimulus that is more likely to cause damage. “But 20%

of people can still detect the pain on the right — which means they have an abnormal ability for pain signals to get through to the brain,” Russo says. The DNIC filter is more likely to fail if a person is stressed, he adds.

Another factor in whether a person develops chronic pain is the initial severity of the acute pain. The first 24 hours after a trauma or operation are thought to be crucial. “If your acute pain is very severe, your risk of chronic pain is much higher,” says Moseley. “If we can reduce the activation of nerve cells that produce the danger message in the spinal cord, then we reduce the chance those nerve cells will sensitize and adapt.” This provides a clue as to how to stop the nerve cells from firing after the injured tissue has healed.

Combine all these ideas, Moseley and Russo agree, and there’s the possibility that doctors can intervene before chronic pain sets in. “If you come up positive on the risk-factor tests, a special rapid-response team will manage your pain in the first 24 hours,” Russo says. These teams would use every pain-killing method at their disposal — from drugs to temporary nerve blocks — to make sure that the patient never rates their pain beyond mild on the pain scale, he adds.

Individuals at risk of developing chronic pain can also be offered targeted psychotherapy to help with the underlying issues that predispose them to it, and to educate them in the mechanisms of pain. In 2014, Toronto General Hospital in Canada became the first centre to implement such a comprehensive programme aimed at preventing chronic postsurgical pain. The team does not have randomized-controlled-trial data yet. But several hundred patients have taken part in the programme, and the results seem promising. “The data suggest we’re doing something right,” says Joel Katz, a pain researcher at the hospital.

NERVE ZAPPERS

For the patients that these early interventions don’t catch — or the millions already living with chronic pain — there are other options in the pipeline, including one that is not a drug at all. Last year, Brooker carried out the first permanent implant of a smart electronic device that stimulates inhibitory neurons in the spine.

The main body of the device is a matchbox-sized titanium box housing all the electronics, which is placed in the fat layer just beneath the skin. A thin wire runs from the device to a metal electrode that is implanted next to the spinal cord.

Brooker’s patient, Jaswir Grewal, had suffered debilitating back pain for decades. After the surgery, he said that the severity of his pain went from eight out of ten to about two or three with the flick of a switch.

A new type of spinal cord stimulation device is in clinical trials.



CHILDBIRTH

Delivering more options for women

Most areas of medicine have changed radically since the 1940s. But women in labour have pretty much the same pain-relief options as their great grandmothers.

For generations, labour wards have offered a trio of escalating pain interventions: a mixture of oxygen and nitrous oxide (gas and air); an injection of the opioid pethidine, which can leave women feeling nauseated and 'out of it'; or an epidural anaesthetic that numbs the lower-body pain, but can restrict the woman to the bed.

But this could be about to change. Last year, midwifery researcher Julie Fleet (pictured) at the University of South Australia in Adelaide and her colleagues conducted a randomized clinical trial that compared pethidine with a nasal spray of the opioid fentanyl.

Fentanyl is not a new drug. But because the body clears it quickly, it was conventionally given through a drip, which restricts movement and limits its appeal on maternity wards.

Around a decade ago, a nasal version of the drug was developed for use by paramedics and on children's wards, where it is now used routinely. Fleet suspected that the reformulated drug could also make a difference in childbirth. The self-administered nasal formulation gives women effective pain relief and allows them to remain mobile during labour.

The researchers showed that although nasal fentanyl and pethidine both controlled pain equally, women who receive fentanyl had shorter labours, less difficulty establishing breastfeeding, and less sedation and nausea. More than 80% of women would use it again, compared with 44% for



pethidine³, says Fleet. "They get the pain relief, but without the sedation, so could feel in control and be active in their labour."

The two hospitals involved in the trial now routinely offer nasal fentanyl to women in labour. Fleet is collecting data to assess whether women who take up this option are less likely to request an epidural.

"There is this big misconception that epidurals are very safe for the baby," Fleet says. "Epidural can be very effective, but it does have increased risks." An epidural is the only pain relief option that requires continual fetal monitoring, because it can cause the mother's blood pressure to drop, which reduces blood flow to the baby and increases the chance that a woman will need a caesarean or an assisted birth. "We think if we give them an option that's less invasive and still effective for pain, they won't need to go on to epidural." **J.M.C.**

Spinal cord stimulation was first trialled in 1967, but it has usually been a treatment of last resort. This is because the simple implants tend to move relative to the spinal cord as the patient moves — even when they breathe. The target nerve is therefore frequently under- or over-stimulated, and neighbouring nerves are hit, too. "You tend to pick up nerves to the ribs, which can be very painful," says Brooker. So people with the implant often turn it down, or even off.

The device that Brooker implanted in Grewal is more sophisticated. Created by start-up company Saluda Medical in Artarmon,

Australia, the device overcomes the problem of electrode movement by continually reading the electrical activity induced in the target nerve, and adjusting its output to keep nerve stimulation within the therapeutic range.

Saluda had already demonstrated the concept's potential using temporary implants, and in October 2015 the company began a multinational three-year clinical trial of permanent devices — which Grewal was part of. While this is taking place, the company is continuing to improve the device, including miniaturizing it. "Making it half as big is not out of the question," says senior vice-president Dan Brounstein.

The Saluda device has impressed pain researchers. "In theory, it's a very significant development," says Russo, whose pain

clinic is participating in the trial. It used to be impossible to know how much of the time the correct level of activation was being delivered to the target nerve. "With this device, it's close to 100% of the time," says Russo.

A wave of similar technologies may be on the way, thanks to an explosion of innovations in spinal cord stimulation. Among the ideas being tested are whether the use of high-frequency electrical impulse patterns suppress pain more effectively, and the use of inductive coupling (the technology behind wireless mobile-phone charging) to power the implant — so that the mobile-phone-sized battery can be worn on the belt rather than implanted under the skin alongside the stimulation device. "It is far more comfortable," says Russo, adding that implanting the device "becomes an outpatient operation".

As the technology has improved, so has the clinical knowledge of which patients will benefit. Those with neuropathic pain from damaged nerves respond the best. "For many years, we were able to achieve 50% of patients achieving 50% pain reduction," Russo says. In the past 4 years, several clinical studies have got close to 75% of patients achieving 75% pain relief. "Once you get to those figures, it no longer makes sense to be a treatment of last resort."

The developments in medication and technology have been welcomed by Michael, who is now 28. He has a spinal implant, and is taking a tailored cocktail of drugs. Together, these therapies have reduced his pain significantly, allowing him to sleep. He has lost 30 kilograms and is mobile, independent, has overseas holidays and an active circle of friends. "Yes he still has pain," Russo says. "But he is living life."

It might be an age-old phenomenon (see page S18), but pain, says Russo, was only established as a medical speciality after the Second World War. "We are the youngest field of medicine," he says, "and changing probably faster than any other."

The fast-blowing winds of change carry the promise of new drugs, devices and early interventions, which many pain clinicians hope will soon translate into better pain-relief options for their patients (see 'Delivering more options for women'). "It's like everything has been thrown up in the air and we're waiting for the dust to settle," says Brooker. "We're waiting to see which of these new toys really is effective once the clinical research is complete." ■

James Mitchell Crow is a freelance science writer based in Melbourne, Australia.

1. Zadina, J. E., Hackler, L., Ge, L.-J. & Kastin, A. J. *Nature* **386**, 499–502 (1997).
2. Zadina, J. E. et al. *Neuropharmacology* **105**, 215–227 (2016).
3. Fleet, J., Belan, I., Jones, M. J., Ullah, S. & Cyna, A. M. *BJOG* **122**, 983–992 (2015).

"If your acute pain is very severe, your risk of chronic pain is much higher."

PERSPECTIVE

JEAN-SEBASTIEN AUSTIN



Equality need not be painful

Pain researchers' arguments for using only male rodents in preclinical pain research don't hold up to scrutiny, says **Jeffrey S. Mogil**.

Sex differences in pain and analgesia are real and robust. I found this out in graduate school, more or less by accident: a casual reanalysis of data from an experiment performed earlier that day revealed that the drug that my colleagues and I were studying worked completely in the male mice but not at all in the females¹. I thought that such a glaring yet unexpected difference would whet my colleagues' appetites for further research. I was mistaken. The postdoctoral fellow I was working with dismissed the finding and encouraged me to get back to work and focus on the "real" phenomenon. Fortunately I ignored him, and I have spent much of my career since studying sex differences in pain, which are as real in people as they are in mice. Yet 25 years later, despite a wealth of evidence and even a change in preclinical-research guidelines, many pain researchers still do not include female animals in their studies.

Women, the data show, are more sensitive to, and less tolerant of, pain than men². However, this quantitative sex difference in pain intensity is much less important than the emerging evidence of qualitative sex differences in pain processing. My colleagues and I have demonstrated, for example, that in the spinal cord, male and female rodents process pain through entirely different immune cells: microglia in male rodents and T cells in females³.

Qualitative sex differences such as these can be seen almost everywhere: stress-induced inhibition of pain, the effects of genes on pain, social modulation of pain and memory of pain. How sensitive a rodent is to pain can even be affected by the sex of the person doing the experiment⁴. These results seem to be the perfect justification for policies that are emerging across the world, including from the US National Institutes of Health⁵ (NIH), that mandate the use of animals of both sexes in preclinical research.

THE 50/50 SOLUTION

Since 2014, the NIH has required the "consideration of sex as a biological variable" in preclinical research. But pain researchers and others still persist with experiments that test only male rodents. Of the 71 research articles that used rodents published in the journal *Pain* in 2015, 56 tested only males, 6 tested only females (4 of these were female-specific studies) and 6 did not disclose the sex. Only 3 papers (4.2%) affirmed the use of both sexes. There is little difference between these percentages and those for the period 1996–2005 (ref. 6).

Why is there so much resistance to using female mice and rats in pain studies? My conversations with many researchers who still use only male animals revealed three major fears.

First, researchers worry that including females in their experiments will increase variability — females have fluctuating gonadal hormone levels — thus necessitating the testing of more subjects. This concern seems superficially plausible, but it is empirically false. Variability in pain data is no higher in female mice than in males⁶ — a fact that has

been shown to be true for biomedical animal research as a whole⁷. This is probably because male animals have their own source of variability: cage dominance hierarchies. Male rodents fight each other for status. Experiments might be affected by which animal is dominant and which is submissive, and by how long ago the aggression occurred.

Second, many researchers think that the NIH policy forces them to double their sample sizes, greatly increasing the cost of experiments. This belief is also false — the policy simply asks for the adequate consideration of sex in all experiments (see go.nature.com/28icfmw). There is no need to use enough animals to ensure that there is sufficient statistical power to detect all quantitative sex differences, which would indeed increase costs, but rather enough to allow the observation of large sex differences, such as those seen in our studies of different immune cells in pain processing³.

This simply requires that half of the subjects are female, a constraint that adds nothing to the cost of the study. Although this 50/50 strategy will not detect small sex differences, it will reveal major discrepancies. It is better to discover such low-hanging fruit than to remain oblivious to all sex differences because no one looked for them. What is more, the discovery of a sex difference can make a paper more interesting, or turn one paper into two. It's hard to see the downside.

A third fear, which I do think is justified, is that reviewers might ask scientists to repeat all their studies in every phase of the oestrous cycle. I would note, however, that there are better ways to study the effects of gonadal hormones⁸ and that many sex differences in pain are actually due to testosterone.

It is crucial not to lose sight of the bigger issue — researchers have an obligation to attempt to solve the problems that are important to society. Most patients with pain are women. We fail in our duties if we conduct research using only male rodents, producing results that might serve only men. The message to my fellow pain researchers is: start including female mice and rats in all your experiments today. You have nothing to lose, and both men and women with pain have everything to gain. ■

Jeffrey S. Mogil is a neuroscientist in the Department of Psychology and Alan Edwards Centre for Research on Pain at McGill University, Montreal, Canada.

e-mail: jeffrey.mogil@mcgill.ca

1. Mogil, J. S., Sternberg, W. F., Kest, B., Marek, P. & Liebeskind, J. C. *Pain* **53**, 17–25 (1993).
2. Mogil, J. S. *Nature Rev. Neurosci.* **13**, 859–866 (2012).
3. Sorge, R. E. et al. *Nature Neurosci.* **18**, 1081–1083 (2015).
4. Sorge, R. E. et al. *Nature Methods* **11**, 629–632 (2014).
5. Clayton, J. A. & Collins, F. S. *Nature* **509**, 282–283 (2014).
6. Mogil, J. S. & Chanda, M. L. *Pain* **117**, 1–5 (2005).
7. Prendergast, B. J., Onishi, K. G. & Zucker, I. *Neurosci. Biobehav. Rev.* **40**, 1–5 (2014).
8. Greenspan, J. D. et al. *Pain* **132**, S26–S45 (2007).

WE FAIL IN OUR
DUTIES IF WE
CONDUCT RESEARCH
USING ONLY
MALE RODENTS,
PRODUCING RESULTS
THAT MIGHT SERVE
ONLY MEN.

Show me where it hurts

Technology for peering into the brain is revealing a pattern of pain, and differences between the acute and chronic forms.

BY SIMON MAKIN

David was lying in the brain scanner, showing no signs of reacting to an intense laser beam shining onto the back of his hand. Several minutes into the procedure, he said: “We should maybe stop this laser.” When asked why, he replied: “It’s starting to feel like when I used to burn my hand with a lighter.”

Thanks to a rare genetic mutation, David cannot feel pain. “Most of us don’t need to think: ‘What does this feel like that I’ve experienced before?’,” says Tim Salomons, a neuroscientist at the University of Reading, UK, who was part of the team running the study that David took part in. “It shows the role pain plays.” Using a common brain-imaging technique called functional magnetic resonance imaging (fMRI) to measure brain activity, the team found that a painful stimulus activated the same regions in David’s brain as it did in healthy controls¹.

Although this seems to cast doubt on the relationship between brain activity and pain, imaging studies in general have revealed much about how the brain processes pain. Some have found patterns that might offer a way to measure pain objectively, whereas others are exploring the differences between acute and chronic pain.

PATTERNS OF PAIN

In contrast with most other senses, conclusive evidence of a brain region dedicated to pain is lacking. Instead, pain is usually associated with activity in numerous areas — a ‘pain matrix’ of regions reliably activated by painful stimuli. These include the somatosensory cortices, which process sensory aspects of pain alongside sensations such as touch and temperature; and the anterior cingulate cortex and insula, which are thought to be important for emotional and motivational dimensions of pain (such as pulling your hand out of a fire). Other areas include the prefrontal cortex (the seat of higher cognitive processes) and the thalamus (a ‘relay hub’ for sensory and motor signals).

Pain-matrix regions are not specific to pain, they are also activated by attention-grabbing stimuli such as flashes of light and loud banging sounds. These stimuli trigger processes involved in detecting important events, directing attention and readying for a response. Because pain also grabs attention, Giandomenico Iannetti — a neuroscientist at University College London who worked with Salomons on the fMRI



Researcher Tim Salomons examines brain scans as a person is subjected to laser stimulation in a scanner.

study — argues that pain-matrix activity may have more to do with the importance of painful events than with the pain itself.

But others think that hidden within that general activity is something more specific. Tor Wager, a neuroscientist at the University of Colorado Boulder, used machine-learning techniques to classify patterns of activity over multiple brain regions — predominantly the pain matrix — to develop a ‘neurological signature’ of pain² (see ‘Signature of hurt’). “We’re showing very specific patterns within these regions that do encode pain,” he says. “Other patterns encode other things, but we can separate them.” For instance, among their other functions, pain-matrix regions are also activated by emotional experiences, such as social rejection and empathy for, or memory of, pain — leading some to say that those feelings also hurt, to some extent. But although rejection and physical pain share a dimension of emotional unpleasantness, heartache is clearly different from being stabbed in the chest — and this distinction can now be discerned in the fine detail of brain images. Wager’s group used its system to distinguish painful heat from non-painful heat; actual pain from anticipation, or recall, of pain; and physical from emotional pain. The group’s algorithm was able to correctly reject the non-pain experiences around 90% of the time, and determine actual pain with more than 90% accuracy — an impressive combination of

specificity and sensitivity. Wager’s system also predicted perceived pain levels and showed that administering a potent opioid drug significantly reduced activation.

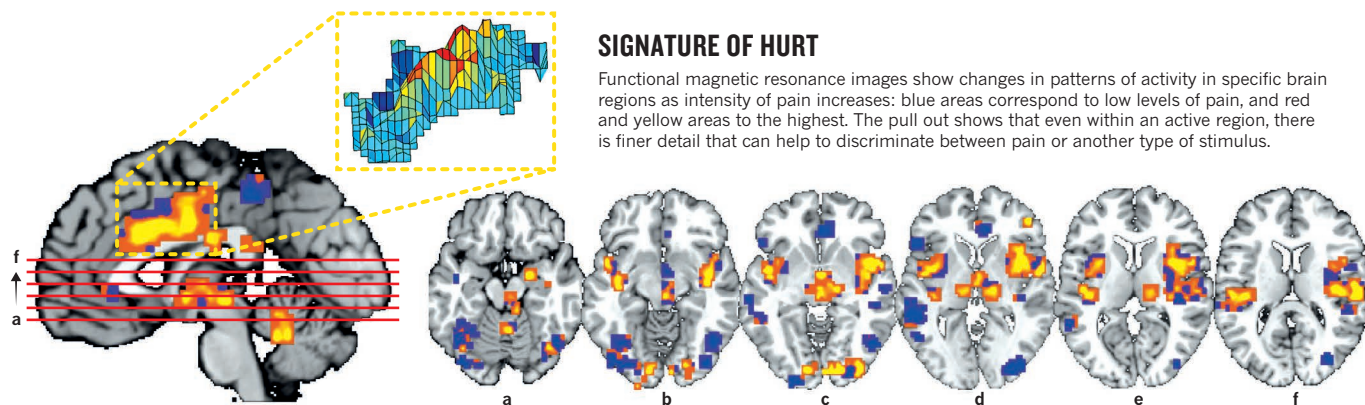
“We’re trying to develop measures that really track the pain that you feel, based on things that come up from the body,” says Wager.

MIND OVER MISERY

Pain can also be influenced by factors such as expectation (which feeds into the placebo effect, see page S14), attention, emotion and even personality. Imaging is allowing researchers to investigate how these elements manifest in the human brain. The ability to exert control through willpower and imagination, known as self-regulation, can alter pain perception. But Wager’s group found that self-regulation had no effect on the neurological signature³. It did, however, affect activity in other brain regions, most notably the nucleus accumbens, which operates through connections to the medial prefrontal cortex to form a circuit within the brain’s reward network. The perception of pain, it seems, is not the result of one system. “The pain signature we developed is a really important component of pain,” Wager says, “but it’s not a complete description.”

Imaging is already helping to determine those other components of pain. Researchers know from animal studies that attention and emotions can modulate pain through a descending system

NATHAN CLARKE



SIGNATURE OF HURT

Functional magnetic resonance images show changes in patterns of activity in specific brain regions as intensity of pain increases: blue areas correspond to low levels of pain, and red and yellow areas to the highest. The pull out shows that even within an active region, there is finer detail that can help to discriminate between pain or another type of stimulus.

TOR WAGER; SOURCE: REF. 2

that connects parts of the brain's cortex and limbic system (the emotion centre) with various regions in the brainstem, which connects with the spinal cord (see page S2). This enables higher brain areas to enhance or inhibit pain signals. "What imaging has proven is when you're sad, anxious or distracted, it doesn't just change the way you express pain, it changes the physiological processing," says neuroscientist Irene Tracey. Her group at the University of Oxford, UK, has been using brain-imaging technologies to examine pain modulation in humans. One study⁴ has already identified which brainstem regions reduce pain signals when a person is distracted, and they are now trying to identify risk factors and brain networks that might make someone more vulnerable to developing chronic pain. "We're using imaging to help explain why someone's painful experience is a particular way — and what mechanisms lock them into that state," says Tracey. "These provide exciting alternative targets for therapies." The hope is that a likely transition from acute to chronic pain can be prevented.

Chronic pain is a huge global burden, affecting around one in five people (see page S4). "We have no scientifically validated treatments for these patients," says physiologist Vania Apkarian at Northwestern University in Chicago, Illinois. "It's a massive health situation." Apkarian's group has found many functional and anatomical features that are unique to the brains of people with chronic pain, helping to establish that acute and chronic pain are fundamentally different. "By definition, that makes it a disease state," says Apkarian. Finding out whether such differences are the cause or consequence of chronic pain is trickier.

To tackle this question, Apkarian's group conducted the first longitudinal brain-imaging study of chronic pain⁵. The researchers followed 39 people with recent back pain for a year, periodically conducting brain scans. Over this period, those who developed chronic pain showed reductions in grey-matter density in the insula and nucleus accumbens. The researchers also found that measures of connectivity between the medial prefrontal cortex and nucleus accumbens taken at the start of the study predicted with around 80% accuracy who would develop chronic pain — stronger

connections conferred higher risk. In a follow-up study, Apkarian's team tracked brain activity associated with perception of back pain and found that, as pain became chronic, activity shifted to brain regions associated with emotion and reward⁶. The extent of this shift was also related to the strength of the connectivity between the medial prefrontal cortex and the nucleus accumbens.

These findings reveal the circuitry that seems to trigger the transition from acute to chronic pain, together with the anatomical changes that are the consequences of it. "This disambiguates the chicken and egg of chronic pain," says Apkarian. His team has also shown that the main determinant of chronic pain is not the injury, but the properties of the person's brain, he adds. And, in rodents, the researchers have been able to block the transition from acute to chronic pain using drugs that inhibit neurons in the nucleus accumbens⁷; a trial is under way to see if this works in humans. "I'm confident we will quickly develop a whole series of new treatment options specific for different types of chronic pain," says Apkarian.

The brain regions that Apkarian identified are the same ones that Wager's group found to be involved in self-regulation. So, although this reward-learning and emotional circuitry is not part of the neurological signature for acute pain, it does seem to play a key part in chronic pain. "We're learning something about how different kinds of pain have different bases in the brain," says Wager. "What's driving your pain might not be the classic pain processes."

"The use of the technology is getting ahead of itself."

PAIN-O-METER

Could these developments bring us closer to being able to measure pain objectively? Such readings would be useful both for drug development and for people who can't express whether they are in pain, such as infants, people in a coma or those with dementia. Several companies in the United States are already offering a service that they say can detect a person's pain signature. And there has been at least one case in which brain scans have been accepted

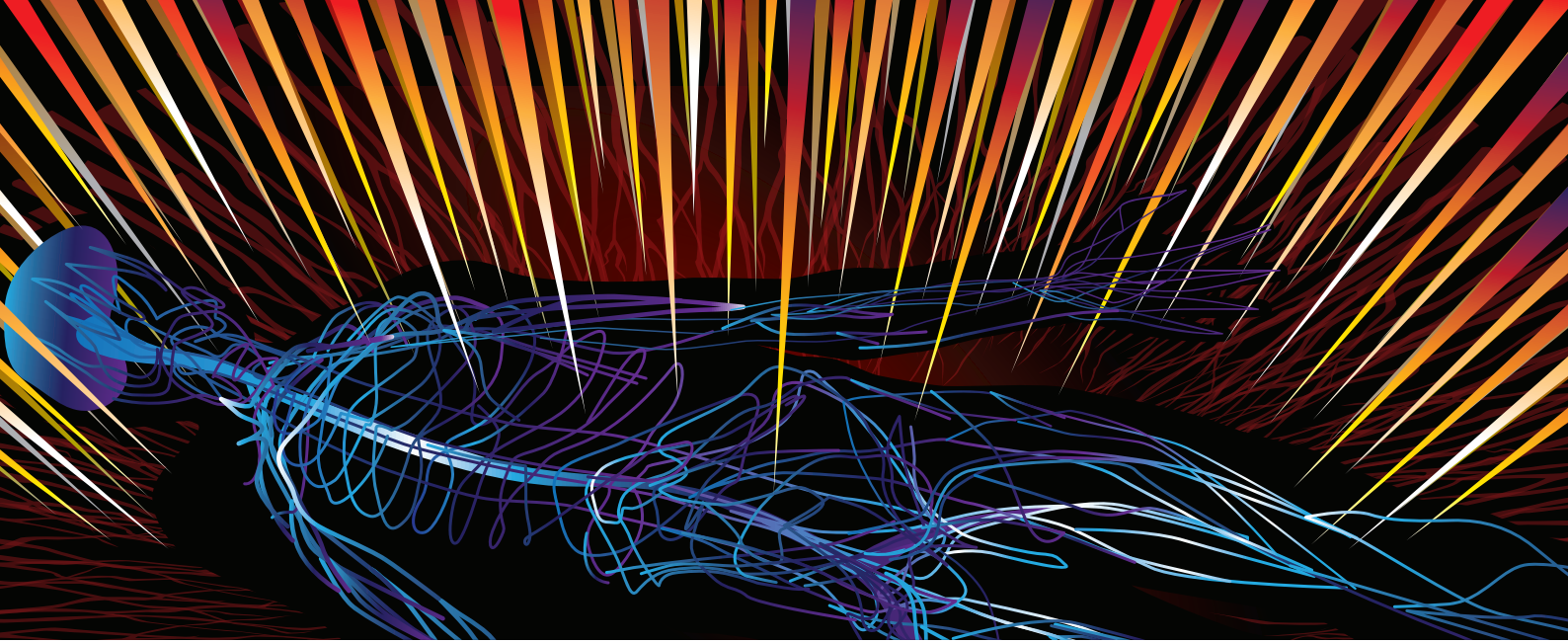
as evidence of chronic pain in US civil courts. But many researchers have grave concerns. Importantly, Wager's results don't apply to chronic pain. "The technologies Tor and others use involve recording how the brain responds to a stimulus," explains neuroscientist Karen Davis of the University of Toronto, Canada. "In chronic pain there's no stimulus, so we need a different approach."

Last December, the International Association for the Study of Pain, based in Washington DC, set up a task force, chaired by Davis, to study the use of brain imaging to identify pain. Over the next year it will produce guidelines on what the technology can and cannot do, whether it is accurate and reliable enough for legal settings, and what the ethical and social issues are. A key concern that Davis and many other researchers have is that fMRI might give misleading results. Certain drugs, for instance, can change vascular function and thus the fMRI signal without having changed brain activity. "Using a vascular-based technology has issues that people haven't been considering," says Davis. Getting this right will be crucial if brain imaging is going to play a part in evaluating pain. "The use of the technology is getting ahead of itself, and there are enormous legal and neuroethical implications," says Davis.

Put someone like David in the brain scanner and you get a false-positive result — he doesn't feel pain even though his pain matrix is active. Conversely, a lack of activity might seem to imply an absence of pain. But most researchers agree that such a conclusion would be unwarranted. "We can confirm pain of certain kinds," says Wager. "But you can never, even in principle, disconfirm pain — because a person's brain might just be unique." ■

Simon Makin is a freelance science writer based in London.

1. Salomons, T. et al. *JAMA Neurol.* <http://dx.doi.org/10.1001/jamaneurol.2016.0653> (2016).
2. Wager, T. D. et al. *N. Engl. J. Med.* **368**, 1388–1397 (2013).
3. Woo, C. W., Roy, M., Buhle, J. T. & Wager, T. D. *PLoS Biol.* **13**, e1002036 (2015).
4. Tracey, I. et al. *J. Neurosci.* **22**, 2748–2752 (2002).
5. Baliki, M. N. et al. *Nature Neurosci.* **15**, 1117–1119 (2012).
6. Hashmi, J. A. et al. *Brain* **136**, 2751–2768 (2013).
7. Ren, W. et al. *Nature Neurosci.* **19**, 220–222 (2016).



NEUROPATHY

A name for their pain

People with neuropathic pain have struggled to find relief with conventional drugs. Researchers are investigating whether more meaningful pain classifications could help.

BY MICHAEL EISENSTEIN

Two years ago, with little fanfare, neurologist Søren Sindrup reported the results of a successful clinical trial¹. On the face of it, it was a modest success story. Instead of coming up with a wonder drug, Sindrup and his team repurposed an existing medication. Nevertheless, some pain researchers consider the trial a potential game-changer — one that marked a turning point in how researchers think about neuropathic pain.

This type of chronic pain arises from damage to the nerves that sense, transmit or process information about environmental stimuli. It can result from numerous initial insults, including spinal cord injury, diabetes and chemotherapy. Patients have generally been grouped on the basis of this initial trauma. But Sindrup, who is at Odense University Hospital in Denmark, and his colleagues took a different approach. They used diagnostic work-ups to cluster patients by their symptoms. This allowed the researchers to home in on a cohort that was more likely to respond to treatment. This is a huge step forward in an area where clinicians have struggled to help their patients. “The drugs we have relieve 50% of pain in somewhere between 1 in 4 and 1 in 7 of the patients we treat,” says Andrew Rice, a pain researcher at Imperial College London. “That’s for the best drugs — and that’s not very good.”

A growing number of pain researchers think that improvements can be found by

analysing symptoms for clues about the underlying nerve damage. Neurologist Giorgio Cruccu of Sapienza University in Rome draws a comparison with another area of neurology. “There is no universal treatment for epilepsy,” he says. Instead, “it depends on the type of seizures”. Pain is a challenging medical target — doctors gain much of their insight from patients’ reports rather than from external observations. But clinicians are attempting to devise more-sophisticated diagnostic tools to give the field a quantitative edge — and perhaps usher this patient population into a new era of evidence-based treatment.

TESTING YOUR PATIENTS

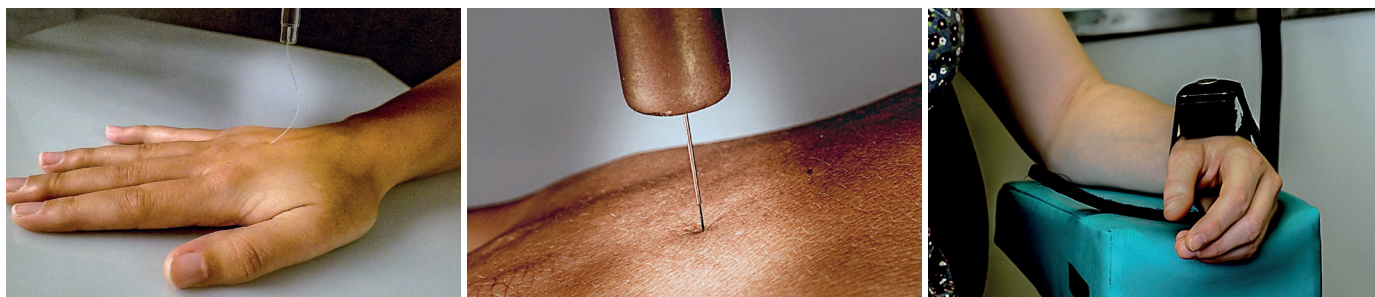
Pain is initially recognized through peripheral sensors in the skin known as nociceptors, which react to potential sources of injury such as heat or mechanical trauma. Nociceptors send signals through specialized nerve fibres to the spinal cord, and from there to the brain (see page S2). Disruption to any part of this process can trigger enduring discomfort, although the severity and sensations experienced — burning or shock-like pain, numbness or tingling — can vary widely depending on the nature of the underlying damage. Not all injuries result in the same pain symptoms. For example, people with post-herpetic neuralgia (which can result after an outbreak of shingles) often have spontaneous pain that resembles an electric shock, but some experience allodynia — pain as a result of benign physical contact, such as clothing rubbing against skin. Over

the past two decades, clinical researchers have come to appreciate that this variety of symptoms offers a way to understand how pain works. “There were hints in the literature that there are different mechanisms at work across various neuropathic pain entities, where patients have the same ‘origin’ of pain, but a different pain mechanism,” says Christoph Maier, a pain specialist at University Hospital Bergmannsheil in Bochum, Germany. “Today, we know this idea is correct.”

If these symptoms do represent different underlying mechanisms, that would help to explain why people in the same patient group respond differently to the same drugs — and that might have implications for treatment. “We have tried to develop a classification that is based on symptoms, which may give some indirect clue about the pain mechanism,” says Nadine Attal, a neurologist at Versailles Saint-Quentin-en-Yvelines University in France. Over the past decade, several questionnaires have been developed, including painDETECT and Douleur Neuropathique 4, which help to distinguish pain associated with nerve injury from that brought on by other causes, and the more detailed Neuropathic Pain Symptom Inventory (NPSI), for further subclassification of patients. These can be completed by patients in minutes, and have proved to be a reliable way to assess the nature and intensity of their pain.

But questionnaires do not objectively measure pain, nor can they zero in on the factors that trigger it. To provide such insights, Maier and other researchers affiliated with the

DANIEL HERTZBERG



Left to right, a whisker-like fibre, pin prick and thermal stimulus are used to test pain sensitivity as part of the quantitative sensory testing protocol.

CHRISTOPH MAIER

German Research Network on Neuropathic Pain (DFNS) have devised a standardized battery of assessments known as quantitative sensory testing (QST). The QST protocol includes components such as hot and cold probes, to determine whether pain is triggered by thermal stimuli, and thin, whisker-like filaments that are applied to the skin to assess sensitivity to touch. “If you have somebody with allodynia, that small filament would feel painful,” says Ian Gilron, an anaesthesiologist at Queen’s University in Kingston, Canada. QST can help researchers to measure the response of different types of sensory nerve, including both the small fibres that detect painful stimuli and the large ones that transmit information about movement and vibration. Although QST enables clinicians to measure and monitor pain symptoms, it is a labour-intensive process that requires extensive training. Furthermore, the variability in pain response across or even within individuals means that QST is better suited to identifying subgroups in a population than for diagnosing individuals.

Skin biopsies taken from the area of pain can provide a more detailed picture of what is happening at the tissue level. “You can demonstrate the loss of small fibres by directly counting how many free nerve endings can be found in the epidermis,” says Cruccu. He also advocates the use of tests that directly measure how well individual nerves function. Such techniques, says Cruccu, “provide objective measures unpolluted by cognitive biases”. Although this type of neurophysiological testing can reveal the nature of nerve damage, it requires costly, specialized equipment and expertise — and some of the more cutting-edge tools have yet to be validated for clinical use.

IN SEARCH OF SUBGROUPS

Researchers are still deciding how to rewrite the diagnostic rule book, but preliminary studies support the idea that a deeper assessment of pain symptoms can lead to more effective care. For example, in Sindrup’s clinical trial¹, although the team recruited patients with diverse neuropathic traumas, it used QST to identify common characteristics that might predict drug efficacy. The researchers found that people with nerves that had become hyper-responsive to temperature or physical probing — the ‘irritable

nociceptor’ phenotype — were more than three times as likely to have pain relief from the anti-convulsant drug oxcarbazepine as those who had the non-irritable phenotype. This response also makes mechanistic sense: Sindrup and colleagues noted that oxcarbazepine blocks the sodium channel proteins that are responsible for nerve signalling, which could well be hyperactive in patients with irritable nociceptors.

This study is one of the few to select patients up front on the basis of pain characteristics, but others have applied similar techniques retrospectively. By using QST and skin-biopsy data collected during a trial of botulinum toxin A, which inhibits the firing of pain nerves, Attal and her colleagues found that people with both allodynia and a higher density of epidermal pain-sensing fibres were more likely to benefit from this treatment². And a team led by Didier Bouhassira, a colleague of Attal’s at Versailles, is preparing to report a study that re-examined data from 1,200 patients who previously participated in unsuccessful clinical trials for a heavily studied neuropathic pain drug. These findings offer hope for improved patient–drug ‘matchmaking’, whereby symptom profiles inform smarter trial design and help doctors to prescribe the treatments that are most likely to be effective.

Integrating data sets from multiple diagnostic approaches offers a way to improve this process. One such effort, by neurologist Roy Freeman at the Beth Israel Deaconess Medical Center in Boston, Massachusetts, and colleagues, analysed QST and NPSI data from past clinical trials to identify four distinct patterns of pain symptoms that seem to correlate in different groups of patients³. These profiles could be developed into ‘fingerprints’ for specific types of neuropathic injury by, for instance, connecting specific pain triggers such as pressure or cold with manifestations of pain such as stabbing or tingling sensations.

Researchers hope that such correlations will reveal information about the roots of pain pathology. A large European patient registry maintained by the DFNS and the public–private organization the Innovative

Medicines Initiative (IMI) is enabling a more thorough hunt for such patterns. “It contains about 4,000 patients,” says Maier, who manages the data set as part of the IMI’s Europain project. “It includes somatosensory profiles, clinical data, QST data, microscopy and skin-biopsy data and, in some cases, genetic data.”

Despite having only a handful of trials to serve as proof of concept, several consortia — including the US-based Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) — are planning on using these phenotyping tools in clinical trials. For now, most of the enthusiasm is coming from the academic sector; pharmaceutical companies expect much stronger evidence before taking on the additional cost. There is also the likelihood that more refined testing will shrink the patient population that drug companies can target with new analgesic drugs. “Instead of getting an approval for all of post-herpetic neuralgia, for example, they’d get one just for post-herpetic neuralgia with allodynia,” Rice says.

Nevertheless, according to Cruccu, a growing number of trials now use the quick questionnaires as a cost-effective fail-safe. Even if, overall, a trial seems unsuccessful, the availability of these data could enable a later search for specific subgroups in which efficacy can be demonstrated. Maier says that findings such as those from Sindrup’s trial suggest that many ‘failures’ may be masking successes: small numbers of patients whose positive response to a drug is drowned out by the sea of people whose pain is poorly matched to the therapy being tested.

For now, the diagnostic tools available give only basic signposts for clinicians who treat people with neuropathic pain. But, given the dearth of effective treatments, even modest gains could have an outsized impact — especially once a next generation of analgesics enters the pipeline. “If there was a way to know who was most likely to respond to a drug and really focus on that in a clinical trial,” says Rice, “that would be magic.” ■

Michael Eisenstein is a freelance science writer based in Philadelphia, Pennsylvania.

1. Demant, D. T. *et al. Pain* **155**, 2263–2273 (2014).
2. Attal, N. *et al. Lancet Neurol.* **15**, 555–565 (2016).
3. Freeman, R., Baron, R., Bouhassira, D., Cabrera, J. & Emir, B. *Pain* **155**, 367–376 (2014).



Researchers such as Anne-Julie Chabot-Doré, pictured in Luda Diatchenko's lab at McGill University in Montreal, Canada, are searching for chronic pain genes.

GENETICS

An incomplete mosaic

Although genetics studies have so far failed to revolutionize pain treatments, some researchers think that a host of discoveries are just around the corner.

BY ERYN BROWN

Neurologist Stephen Waxman wants to understand how genes influence chronic pain. He hopes that unraveling the mystery will offer relief to the patients he studies. Some are in agony because of old injuries, others because of diabetic nerve damage. And there are those who battle rare disorders such as inherited erythromelalgia, experiencing searing pain in their extremities when they come into contact with mild warmth or engage in moderate exercise. “For these people, putting on socks is like having hot lava poured on their bodies,” Waxman says.

Researchers such as Waxman, who is at Yale University's Center for Neuroscience and Regeneration Research in West Haven, Connecticut, are betting that analysing patients' DNA will help to explain the underlying causes and mechanics of chronic pain, which now afflicts around one billion people worldwide. Ultimately, researchers hope that such explorations will lead to better treatments for those who have chronic pain, by revealing targets for a new generation of drugs that are capable of targeting pain without dulling the

senses or promoting addiction.

But this is a search that has already gone on for a couple of decades. And although scientists have discovered that genetics have a significant role in pain — anywhere from around 20% to 60% of the variability in how people experience pain is attributable to differences in genes — no one has yet pinpointed any crucial smoking guns in DNA. And the complexity of the problem can be a bit depressing. “Once you realize something is mediated by, say, 1,000 genes, you wonder if it's even possible to figure it out,” says Jeffrey Mogil, a pain researcher at McGill University in Montreal, Canada. But other researchers are more optimistic, and suggest that continuing with the approach, with some refinements, will yield useful discoveries.

ELEMENTAL CONCERNS

Across the campus from Mogil's office, molecular geneticist Luda Diatchenko is undeterred in her search for pain genes. Diatchenko thinks that the large number of unknowns means that scientists are actually on the verge of a flood of discoveries. It took time, she says, to design pain-genetics experiments and to develop methods for studying the genome that

were not prohibitively expensive. “The studies have not been done yet,” Diatchenko says. “It will be an explosion — soon.”

Pain-genetics researchers have pursued two main avenues of inquiry, says neuroscientist Stephen McMahon at King's College London. The first strategy, taken by researchers such as Waxman, is to study rare pain disorders that run in families to identify single-gene mutations. This approach has produced a handful of tantalizing leads. For instance, inherited erythromelalgia, which Waxman says affects around 30–40 families worldwide, is caused by a mutation in a gene that causes the sodium-ion channel $Na_v1.7$ to become overactive. This protein is crucial for conducting pain signals in peripheral nerves, but not, it is thought, in the central nervous system.

If drug developers could use this information to come up with a way to dial down $Na_v1.7$ in people with chronic pain, they could develop systemic treatments that would dampen pain signals in nerve cells without causing side effects such as sleepiness, confusion, loss of balance or addiction, says Waxman. Drug companies such as Amgen, Pfizer and Convergence Pharmaceuticals are

LUDA DIATCHENKO

working on $\text{Na}_v1.7$ inhibitors, but Waxman doesn't expect to see any therapies approved for several years. "Finding a drug that's selective for this sodium channel is tricky," says McMahon. And accidentally hitting another subtype of sodium channel — such as the ones that are essential for controlling heart-beat — would be dangerous.

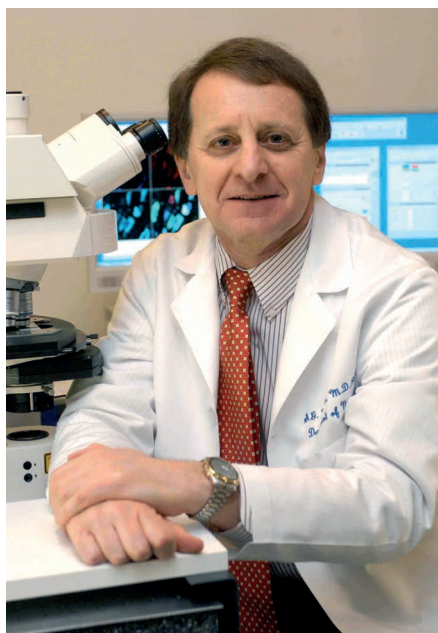
The second approach is broader: rather than just looking at rare inherited conditions, the DNA from large cohorts of patients is sequenced to try to identify genetic variants and the traits, or phenotypes, that correlate with them. The hope was that such studies would reveal a small number of key pain genes — those shared by all or many people with various chronic pain disorders. But what researchers found instead was that pain, like many chronic conditions, is caused by a complex interaction between genes and the environment, influenced by hundreds, if not thousands, of genes in each individual.

William Maixner, director of the Center for Translational Pain Medicine at Duke University in Durham, North Carolina, refers to the causes of chronic pain as "a mosaic of pathways" within each individual that change over time owing to environmental factors, and that affect psychological processes as well as those related to nerve damage. Maixner is working with Diatchenko and others to tease out the genetic mechanisms that are at work in a number of pain disorders, including irritable bowel syndrome, fibromyalgia and lower back pain.

Their Orofacial Pain: Prospective Evaluation and Risk Assessment, or OPPERA, study focuses on people with temporomandibular disorder (TMD), a common facial-pain condition of unclear origin. The team collected data for up to 5 years from just over 2,700 TMD-free men and women, 260 of whom developed TMD during the study. Maixner's group examined DNA variations known as single nucleotide polymorphisms (SNPs) in 358 genes that regulate pain, and tracked 202 phenotypes in the volunteers¹. The initial findings underscored the complexity of TMD. The researchers failed to identify any single genetic variation associated with the condition, but did find five SNPs linked with risk factors for TMD, including ones related to non-specific facial pain, physical symptoms, stress and negative mood. The team has since expanded its analysis to the entire genome of the OPPERA participants, and hopes to publish updated results by 2017.

Researchers such as Mogil and McMahon, who are vexed by the slow progress in pain genomics, say that genome-wide association studies have been too small to detect culprit variants — and that the funding isn't there to

"The studies have not been done yet. It will be an explosion — soon."



Neurologist Stephen Waxman studies the genetics of rare pain disorders.

support larger-scale efforts.

Compounding the problem is the difficulty of correctly phenotyping people with chronic pain. Christopher Sivert Nielsen, a pain psychologist at the Norwegian Institute of Public Health in Oslo, says that pain disorders that receive separate diagnoses are often not that different from each other. Furthermore, chronic pain is very common, making the identification of controls challenging. In this sense, chronic pain is very different from well-characterized diseases such as multiple sclerosis (MS). "If you study MS, you go into the clinic, you collect cases and the rest of the world provides your controls," Nielsen says. "But this doesn't work for pain." Because many types of pain have common genetic origins, it's easy for a person with a related pain type to end up in a control group, disrupting an association study. To pin down phenotypes, Nielsen adds, researchers will need to screen study participants much more rigorously — a difficult task in the large groups required to do genome-wide association studies well.

Maixner says that OPPERA researchers are trying to cut through the noise in their data by devising new methodologies to understand how genes relate to symptoms. Bioinformaticians are working on stratification procedures that divide the study population into three distinct subgroups — a pain-sensitive cluster of people with heightened sensitivity to experimental pain stimuli, a global-symptoms cluster with pain sensitivity and psychological distress, and a third group with neither². Most people with TMD fall into the first and second groups. By analysing these clusters instead of the entire population, Maixner suggests that scientists will find it easier to tease out how

genes contribute to the development and to the manifestation of symptoms in each group.

Some researchers are looking beyond the genome to the epigenome — the markers on DNA that have been added by processes such as methylation. Epigenetic changes alter gene expression and therefore affect various biochemical events. "The hope is to discover a whole new number of cellular processes that control the dynamics of a chronic pain state — the 'on' switches," says McMahon.

He says that studying epigenetics could help scientists to understand the environmental influences that make a person more likely to develop chronic pain. He thinks it could also improve treatments, reasoning that drugs that interfere higher up the chain of biological events should prove more effective than therapies that operate at the periphery. "Using Na_v blockers is like attacking the foot soldiers, whereas disrupting epigenetic processes is like taking out a general giving an order to the whole army," he says.

In 2014, McMahon collaborated on a study that looked at identical twins with different levels of pain sensitivity. The team found³ methylation differences connected to several genes, including the pain gene *TRPA1*. But epigenetics is dizzyingly complex, and other researchers note that it is not yet possible to link epigenetic changes to the environmental factors that might have caused them. "It's early days," concedes McMahon, who stresses that merely finding epigenetic or genetic targets isn't enough — such work must be accompanied by experiments in cells and model organisms that explain the biology going on in the cells.

And Mogil wonders if "epigenetics is a bit of a flavour of the month". Genetics failed, so researchers are now asking "Where will we bet all our chips now?" he says. For now, Mogil has largely shifted the focus of his research from identifying pain genes to validating known candidates in rodent studies, and to understanding how sex impacts pain processing (see page S7). This is the type of follow-up work that could help researchers to understand the fundamental mechanisms of pain and what the genetic and epigenetic findings have to do with them.

Researchers are not yet close to understanding the genetic components of pain well enough to produce tailored pain therapies to satisfy the push for precision medicine, says Mogil. But he is not pessimistic — just realistic. There is still a lot to learn, and for a researcher that is good news. "What I love about genetics is, you can find interesting proteins to study without even knowing what you're looking for." ■

Eryn Brown is a freelance writer based in Los Angeles, California.

1. Slade, G. D. et al. *J. Pain* **14**, T116–T124 (2013).
2. Bair, E. et al. *Pain* **157**, 1266–1278 (2016).
3. Bell, J. T. et al. *Nature Commun.* **5**, 2978 (2014).



PLACEBOS

Honest fakery

Armed with a clearer understanding of how placebos work, researchers are suggesting that inactive substances might be used to mitigate chronic pain.

BY JO MARCHANT

In April, Ted Kaptchuk addressed hundreds of physicians and scientists at the Behind and Beyond the Brain symposium in Porto, Portugal. Within minutes, ripples of laughter were spreading around the conference hall.

Kaptchuk, a researcher at Harvard Medical School in Boston, Massachusetts, was showing the audience a cartoon in which a doctor hands over a prescription note. “I want you to take this placebo,” says the white-coated medic to her bemused patient. “If your condition doesn’t improve, I’ll give you a stronger one.” The chuckles were a response to the absurdity of openly treating a patient with fake pills. By definition, placebos have no active ingredient, so the idea that someone might benefit from knowingly taking one — let alone that different placebos could have different effects — seems nonsensical. But Kaptchuk invited his audience to take the scene seriously. Honest placebos can work, he insisted. And some placebos really are stronger than others.

Kaptchuk’s trials are overturning many assumptions about the best way to care for

patients, particularly those in pain. After four decades of probing the mechanisms of placebo responses, researchers are advancing the argument that inert pills are more than just negative controls in clinical trials: they can be a treatment in their own right.

PLEASING MEDICINE

The modern idea of the placebo effect stems from 1955, when US physician Henry Beecher analysed the results of 15 studies and concluded that, regardless of a patient’s complaint, around one-third showed a significant response to a placebo¹. The effect is now well-established, particularly for conditions that rely on subjective reports, such as pain.

There are lots of reasons why someone in a clinical trial might feel better. Symptoms often ease with time, or trial participants might report an improvement to please the experimenters. Because of this, placebo responses are commonly viewed as illusory — a baseline against which to compare the action of new drugs. But there is now a large body of research showing that the effects of placebos can be very real.

Fabrizio Benedetti, a placebo researcher at

the University of Turin, Italy, points to a 1978 study² by neuroscientist Jon Levine that, he says, represents the moment that “the biology of placebo was born”. Levine and his colleagues administered intravenous infusions of saline to patients who were recovering from surgery, telling them that it might be morphine. One-third of them reported a significant reduction in pain. Then, the researchers secretly added naloxone, which blocks the action of painkillers such as morphine by binding to opioid receptors in the brain, to the infusions and the patients’ pain returned. Levine had shown that a placebo response could be biochemically blocked.

Levine’s study was revolutionary because it suggested that patients don’t simply imagine or pretend that their pain is eased with placebos. Their analgesia reflects a measurable, physical change — mediated by the release in the brain of endogenous opioids called endorphins². This finding has since been confirmed by dozens of brain-imaging studies, which show increased binding of endorphins to opioid receptors in response to placebo painkillers, as well as reduced activity in areas of the brain involved in processing pain³.

DANIEL HERTZBERG

Endorphins aren't the only neurotransmitters involved. Placebos can activate endocannabinoids (which bind to the same receptors as the psychoactive constituents of cannabis) or dopamine, or they can reduce the levels of prostaglandins (which dilate blood vessels and increase sensitivity to pain). In general, Benedetti says, "placebos can modulate the same biochemical pathways that are modulated by drugs".

Inert substances cannot, of course, create biological changes. A placebo's active ingredient, says Kaptchuk, is a person's psychological response to being treated. Tor Wager, a neuroscientist at the University of Colorado Boulder, agrees. His functional magnetic resonance imaging (fMRI) studies were among the first to show that placebos reduce activity in relevant brain areas when people are subjected to pain. But before the onset of pain, his fMRI scans show something different: receiving a placebo increases activity in the two parts of the brain involved in emotion and valuation, the prefrontal cortex and the ventral striatum³. "We think the placebo is causing a re-evaluation of the pain," concludes Wager. "It doesn't mean the same thing to you."

LEARNING NOTHING

Placebos influence expectation: how good or bad we think our pain is going to be. This expectation is influenced by what we're told about a treatment and also its nature — invasive treatments (such as surgery or acupuncture) often elicit larger placebo responses than interventions that seem more modest (such as pills). Social factors including the attitude of the practitioner can also influence patients' symptoms^{4,5}. What's now coming to light, however, is that placebo responses can also be learned. Just as Russian physiologist Ivan Pavlov discovered that dogs salivate in response to a buzzer associated with food, similar mechanisms are thought to drive placebo responses previously assumed to rely purely on conscious expectation.

For example, giving volunteers several doses of a real painkiller — or surreptitiously reducing the strength of experimental pain — makes subsequent placebo responses to the same stimulus stronger and more consistent. Benedetti calls this process "pre-conditioning". When he and neuroscientist Luana Colloca, now at the University of Maryland in Baltimore, subjected volunteers to electric shocks, pre-conditioning resulted in a five-fold boost to the average pain relief conferred by a placebo⁶.

In some circumstances, such learned responses can override conscious expectations. Wager and his colleagues reported that after four episodes of pre-conditioning, an inert cream reduced pain in volunteers even when they knew it was a placebo⁷. "Eventually, it doesn't matter what you think, because

your brain has learned," says Wager.

Different drug memories can trigger different neurochemical pathways. Benedetti demonstrated this effect by pre-conditioning some volunteers with morphine and others with the non-opioid painkiller ketorolac⁸. The subsequent placebo response of those in the morphine group involved endorphin release, whereas in the ketorolac group it was mediated by endocannabinoids. "It shows that not all placebos are equal," says Benedetti.

The key question is whether these drug-like placebo responses can be harnessed in medical care. Patients could benefit from measures such as using language designed to boost expectations or to strengthen the social bond between doctor and patient⁴. But researchers are now suggesting something previously unthinkable — a role for placebos themselves.

Colloca suggests that, by taking advantage of learning mechanisms, doctors could give placebos honestly and reduce the amount of medication. For example, a doctor might pre-

"The assumption has been that deception or concealment is necessary for placebos to work."

scribe a blister pack of painkillers, and tell the patient that it contains both drugs and placebos — but not which pills are which. Earlier this year, Colloca and her colleagues reviewed 22 studies that used similar techniques, covering conditions such as insomnia, autoimmune diseases and pain⁹. They concluded that these approaches have the potential to reduce side effects (although some of these may be conditioned responses, too), limit problems with drug dependency and toxicity, and reduce costs.

Benedetti loves the idea. "This is one of best applications of placebos in clinical practice," he says. In a trial published in February, he showed that in people with Parkinson's disease, pre-conditioning with the drug apomorphine made patients respond to a placebo just as strongly as they did to the active drug¹⁰. Alternating drugs and placebos might delay the development of tolerance, he suggests.

Kaptchuk is going one step further. For conditions such as chronic pain, for which placebo effects are large, drugs aren't very effective and taking them can have downsides (see page S4), he suggests sometimes ditching medication altogether and openly giving placebos. He made headlines in 2010 with a placebo study for irritable bowel syndrome (IBS) in which patients were told that they were receiving a sugar pill¹¹. "Historically, the assumption has been that deception or concealment is necessary for placebos to work," Kaptchuk says. "My logic was that maybe we could tell patients upfront that placebos may work and tell them to give it a try." The results were startling: 59% of patients who knowingly took sugar pills reported adequate relief from their symptoms, compared with 35% in the no-treatment group — better than most

IBS drugs, he adds. "I was very surprised by the results," says Kaptchuk, "even though I hoped it would work."

And it wasn't a fluke. At the symposium in Porto, Kaptchuk followed the cartoon with the results of a new test of an open-label placebo. The trial included 97 patients with chronic lower back pain who had not responded to previous therapies. All continued their usual treatment, but those randomized to the open-label placebo group were also given twice-daily sugar pills, along with an explanation of the research behind why these might help them.

Over three weeks, patients in the placebo group reported a marked drop in pain, whereas the pain of the treatment-as-usual group didn't significantly change. The open-label placebo triggered "sometimes modest, sometimes dramatic, improvements in pain and disability that had major impacts on people's lives," says lead researcher Cláudia Carvalho, a psychologist at the ISPA-University Institute in Lisbon.

Carvalho and her co-authors are still not sure why placebos seem to help patients who haven't responded to treatments in the past. Carvalho suspects that for some, knowingly taking placebos may have made them more aware of the role of the mind in controlling pain. "It empowered them and changed their relationship with their pain," she says.

More studies of honest placebos are in the pipeline — other teams are conducting trials in cancer-related fatigue and depression, and Kaptchuk is recruiting for a trial that aims to replicate and extend his original findings in IBS. If the results continue to be positive, Kaptchuk suggests that for appropriate conditions, placebos — honestly prescribed by clinicians — could become a routine part of medical care. "Placebos have always been a negative for medicine," he says, "but for many patients, trying open-label placebos could be a first line of treatment before any drugs are prescribed." ■

Jo Marchant is a freelance science journalist based in London, and author of *Cure: A Journey into the Science of Mind Over Body* (Canongate, 2016).

1. Beecher, H. J. *Am. Med. Assoc.* **159**, 1602–1606 (1955).
2. Levine, J. D., Gordon, N. C. & Fields, H. L. **312**, 654–657 (1978).
3. Wager, T. D. & Atlas, L. Y. *Nature Rev. Neurosci.* **16**, 403–418 (2015).
4. Finniss, D. G., Kaptchuk, T. J., Miller, F. & Benedetti, F. *Lancet* **375**, 686–695 (2010).
5. Moerman, D. & Jonas, W. B. *Ann. Intern. Med.* **136**, 471–476 (2002).
6. Colloca, L. & Benedetti, F. *Pain* **124**, 126–133 (2006).
7. Schafer, S. M., Colloca, L. & Wager, T. D. *J. Pain* **16**, 412–420 (2015).
8. Amanzio, M. & Benedetti, F. *J. Neurosci.* **19**, 484–494 (1999).
9. Colloca, L., Enck, P. & DeGrazia, D. *Pain* <http://dx.doi.org/10.1097/j.pain.0000000000000566> (2016).
10. Benedetti, F. et al. *J. Physiol.* <http://dx.doi.org/10.1113/JP271322> (2016).
11. Kaptchuk, T. J. et al. *PLoS ONE* **5**, e15591 (2010).



M. R. Rajagopal, founder of palliative-care charity Pallium India, attends to a patient in Thiruvananthapuram, India.

PALLIATIVE CARE

The other opioid issue

The ‘war on drugs’ has left many in developing nations with no access to strong painkillers. But governments in these countries are learning to stop worrying and love the poppy.

BY LUCAS LAURSEN

When pharmacologist Ravindra Ghooi learned in 1996 that his mother had terminal breast cancer, he began to investigate whether he could obtain morphine, in case she needed pain relief at the end of her life. But a morphine prescription in India at that time, even for the dying, was a rare thing: most states required four or five different licences to buy painkillers such as morphine, and there were harsh penalties for minor administrative errors. Few pharmacies stocked opioids and it was a rare doctor who held the necessary paperwork to prescribe them. Ghooi, who is now a consultant at Cipla Palliative Care and Training Centre in Pune, used his connections to ask government and industry officials if there was a straightforward way of obtaining morphine for his mother. “Everybody agreed to give me morphine,” he recalls, “but they said they’d give it to me illegally.”

Jim Cleary, an oncologist and palliative-care specialist at the University of Wisconsin–Madison, has heard similar stories. “Patients with pain have been unwitting victims of the war on drugs,” he says. Opioids have been a hot potato since the 1961 United Nations Single

Convention on Narcotic Drugs. The US-led war on drugs that followed resulted in widespread reluctance to prescribe and supply opioids for fear that patients would become addicted or overdose, or that drug cartels would divert opioids to the black market. Cleary says that countries such as the United States have an “unbalanced” opioid situation, and that abuse in these countries has distorted policies elsewhere, restricting legitimate access.

More than 5 billion people worldwide cannot get the medical opioids that they need. That is a staggering amount of unnecessary agony. “It’s about the government and society at large accepting that while we have a responsibility to prevent abuse and diversion of opiates, we also have a responsibility to people in pain,” says former anaesthesiologist M. R. Rajagopal, who is a founder of palliative-care charity Pallium India in Thiruvananthapuram.

Cleary has been working with activists such as Ghooi and Rajagopal, advocating policies that recognize pain relief as a human right. Such work is slowly reshaping some national drug policies, but changing the status quo entails more than just arguing that patients have a right to opioids — campaigners must

also help governments around the world to chart a course to the safe supply of these powerful painkillers.

ALL SORTS OF BARRIERS

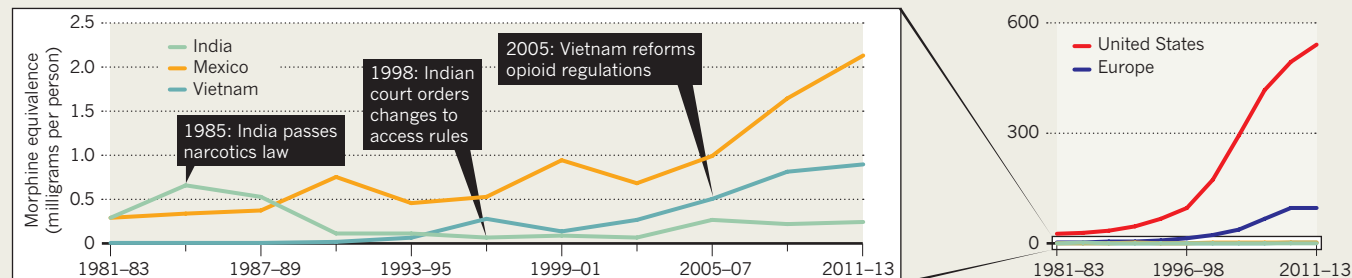
Many countries have reason to be cautious. In Vietnam, French colonists once promoted opium use to create dependency and enable control of the population. Vietnam now has the highest rate of injection drug use in southeast Asia. In Mexico, the illegal narcotics trade has caused violence, and a level of displacement that approaches that of war zones. It is not surprising, therefore, that people in many countries associate opioids with moral issues, not just medical use.

India’s law governing narcotic drugs was introduced in 1985. It was so strict that there were serious consequences for medical staff or private carers who made errors in record-keeping. Fear of addiction grew — even for people who were dying. By 1998, India’s morphine consumption had dropped by 92%. That year, Ghooi filed a public-interest litigation in the High Court of Delhi, calling for the procedure for the supply of morphine to be simplified. The court eventually ordered Indian states to speed up their opioid-licensing processes.

TOM PRICE

OVER AND UNDER

After decades of cultural, legal and logistical barriers, many developing countries are improving access to opioids, but use is still much lower than in developed countries. The United States is one of the biggest consumers of the drugs, mainly owing to the popularity of oxycodone.



The morphine-equivalence metric represents five opioids used to treat moderate to severe pain. Data show amount distributed to dispensaries, rather than individual consumption.

The ruling in Delhi was only the start. Its impact depended on how fast the 25 states that India had at that time would comply. By 2006, only ten had updated their opioid rules. The following year, Ghooi, Rajagopal and Poonam Bagai — founder of Delhi-based children's cancer charity CanKids...Kids-Can — petitioned India's Supreme Court, demanding that the central government do more to force the process.

Cleary's Pain and Policy Studies Group (PPSG) at the University of Wisconsin-Madison is trying another approach. It runs an international fellowship programme that encourages experts from low- and middle-income countries to tackle barriers to opioid use in their own country. In 2012, three Indian palliative-care experts joined the programme, and on return to India began meeting with government officials to discuss a strategy for better opioid access. Working alongside Pallium India, they assisted in drafting an amendment to the law, which was passed in 2014, streamlining opioid permissions to just one licence and shifting powers over drug policy from state to central government.

Cultural barriers are harder to overcome. A 2014 study found that nurses in India rarely recommended prescribing painkillers, even when they thought that the patients needed them, and doctors rarely asked whether they were needed (V. LeBaron *et al. Oncologist* **19**, 515-522; 2014). The study authors described a "culture of non-intervention".

When it comes to bringing about policy change, the difference between failure and success can depend on the decision of a single government official — and activists must have patience. "I now realize that things change slowly," Rajagopal says. "Usually, officials change, so I can bide my time and try again." Nevertheless, Rajagopal is seeing progress. "Good experiences, once made visible, do tend to get replicated elsewhere," he says. Cities and states are more likely to increase access to opioids and palliative care when they see positive

results, without an increase in drug-related problems, in neighbouring regions. By April this year, the situation had changed so much that the Supreme Court of India closed the 2007 case, writing, "the present petition does appear to have served its purpose".

DIFFERENT MODELS

The New York-based organization Human Rights Watch has used a similar advocacy approach in its work to reform Mexico's prescription-opioid regulations. But interim director for health Diederik Lohman says that this is not the only way to change policy. "The other approach is more of an inside game," he says, and it involves helping ministries and drug regulators to hire palliative- and pain-care specialists who can then work within the organization's system to offer an international standard of care.

The international non-governmental organization the Global Access to Pain Relief Initiative (GAPRI) has taken this inside approach in Nigeria by supporting a full-time fellow to assist the country's director of food and drug services with morphine procurement, distribution and usage tracking. This strategy means that "ownership is with the ministry", Lohman says — the government has a commitment to the programme, making it more likely to succeed.

Many palliative-care advocates end up using both strategies. In addition to PPSG's advocacy work in India, in other countries it has taken an inside-government approach. In 2005, PPSG participated in a palliative-care workshop led by the Vietnamese Ministry of Health. Two of the ministry officials were then invited to complete year-long PPSG fellowships. The officials translated Vietnamese laws surrounding opioid accessibility into English, compared them with World Health Organization guidelines, and drafted new regulations for the Ministry of Health to consider. Since then, opioid consumption in the country has risen nearly 3-fold, from 0.3 mg per person in 2005 to 0.9 mg in 2013 (consumption in the United States in 2013 was more than 500 times higher at 520 mg

per person; see 'Over and under'). Some government officials were already open to reforming the morphine regulations, but the fellowship gave them the opportunity to gather the information they needed to help persuade their colleagues.

FINDING CONSENSUS

In April, the UN General Assembly held a special session on the world drug problem, which reflected the international community's slow tilt towards balancing the benefits and risks of opioids. Many countries still want a "drug-free world", but a growing number have concluded that the war on drugs does not work, Lohman says, and these countries are taking a more public-health-centered approach. Despite the two sides still differing on their drug-control position, the assembly came to a consensus to ensure the availability of controlled substances for medical use. "In that respect, we've made a giant leap forward," Lohman says.

But terminally ill people without access to pain relief may not consider this leap big enough. The UN statement does not constitute legislation, and even in countries, such as India, that have reformed their laws, "legislation does not necessarily mean practice", Rajagopal says.

Ghooi found that out the hard way when, in 2010, his wife and fellow pharmacologist Shaila also developed cancer. The disease quickly became terminal, but Shaila's cultural suspicion of opioids meant that even though she was in incredible pain and morphine was readily available, she resisted taking the drug. "My wife grew up in that generation," recalls Ghooi. She thought that "morphine was too strong an analgesic and she was not ready to take it", he says. But he thinks that change is coming. "I do feel hopeful that, in five or ten years, at the rate this government is going, things will be better."

But Rajagopal, who spent 19 years trying to get India's main narcotics law changed, doesn't see an end coming soon: "I don't see any alternative to slow, painstaking work." ■

Lucas Laursen is a freelance journalist based in Madrid.



Painful progress

~2000 BC

Mesopotamians and Egyptians recognize different types of pain, such as burning and stabbing. Where there is no obvious injury, pain is attributed to demons, ghosts or gods.

Mesopotamians think that these attack by touching or striking the body; Egyptians say that the spirits enter the body through the ears and nostrils. Physicians sometimes use narcotics such as opium or the poisonous plant henbane (*Hyoscyamus niger*) to relieve pain, but treatment consists mainly of spells or prayers.

~410 BC

Greek physician Hippocrates and his followers dismiss supernatural causes of pain, arguing that it is a symptom of disease. Hippocratic medicine considers pain to be a useful clue to what is wrong with the patient. Among the texts known as the Hippocratic Corpus are instructions in the art of diagnosis. High on the list of questions that physicians were to ask patients are the familiar "Are you in pain?" and "Where does it hurt?"

1664

French philosopher René Descartes proposes specific pain pathways from the point of origin to the brain; the concept holds for 300 years. He illustrates the mechanism in his book *Treatise of Man* — a boy with his foot near a flame is hit by "particles of fire", which speed along a nerve to the spinal cord and on to the brain, where his soul lies. The soul transforms the signals to a perception of pain, releasing "animal spirits" that course through the nerves to the leg, prompting it to move. The book is published posthumously so that Descartes can avoid the wrath of the Church, which teaches that pain is a gift from God.

1798

English chemist Humphry Davy tests the effects of inhaling nitrous oxide. It makes him giggly and dizzy, but also eases the agony of an erupting wisdom tooth. "The pain always diminished after the first four or five inspirations," he says. Later, Davy reports how a mix of nitrous oxide and oxygen produces reversible unconsciousness in animals. He suggests that the gas "may probably be used with advantage during surgical operations", although the idea of gaseous anaesthesia languishes until the 1840s.



1805

German pharmacist Friedrich Sertürner isolates morphine, the active ingredient of opium. The milky gum tapped from unripe seed pods of poppies (pictured) had been used to deaden pain since prehistoric times, but despite improved preparations (such as laudanum) the variable potency of plant products made their effects unpredictable. Morphine proves ten times as potent and more reliable than opium, making it a mainstay of pain relief.



-6000 BC LEAF RELIEF

People of the Nanchoc Valley in Peru are the earliest known users of cocaine. Archaeological evidence suggests that they chewed coca leaves (pictured) with quicklime to speed the release of the drug — a traditional painkiller in parts of South America.

-2250 BC EARLY WORMS

A Babylonian clay tablet advises treating the pain of a burrowing 'tooth worm', which is thought to be the cause of caries, by plugging the hole with gum mastic and powdered henbane seeds. This is the earliest known written prescription for a painkiller.

47 BRIGHT SPARK

Roman physician Scribonius Largus prescribes electrotherapy for headaches and gout. In his medical text *Compositiones*, he recommends placing the electric ray *Torpedo marmorata* on the brow or under the feet, allowing it to discharge its electricity "until the patient's senses were benumbed".

1898 TRIAL BY ORDEAL

German surgeon August Bier proves the effectiveness of spinal anaesthesia. He administers cocaine to his assistant through a lumbar puncture, then burns and hammers the assistant's legs, finishing by twisting and squashing his testicles. The assistant feels nothing — until the anaesthetic wears off.

For thousands of years people have sought explanations for pain and ways to ease it. Despite a better understanding of the mechanisms behind the sensation, much remains baffling, and the search for better treatments continues. By **Stephanie Pain**

1899 BARK WITH BITE

German company Bayer creates aspirin. The drug has its origins in an age-old remedy for aches and pains — willow bark. The powdered bark contains the analgesic salicin, which Bayer modifies to create the less-toxic acetylsalicylic acid. Aspirin is now one of the world's most widely used drugs.



1906 ALARM SYSTEM

British neurophysiologist Charles Sherrington proposes the existence of nociceptors — specialized nerves that detect potentially harmful stimuli, such as extreme temperature. If the intensity is enough to cause injury, the nerves relay a pain signal to the brain.

1936 RIGHT JAB

Anaesthesiologist Emery Rovenstine establishes the first pain clinic at New York City's Bellevue Hospital, where he pioneers new methods for nerve blocking. Injections of anaesthetic into nerves ease the pain of angina, sciatica, neuralgia and some cancers.

2004 BRAIN DRAIN

People with chronic back pain are shown to lose as much as 11% of their brain tissue (A. V. Apkarian *et al.* *J. Neurosci.* **24**, 10410–10415; 2004). Subsequent studies find that other causes of chronic pain, such as persistent headaches and irritable bowel syndrome, also lead to shrinkage of grey matter.



2014 BACK TO THE FUTURE

An analgesic is discovered in the herb *Corydalis yanhusuo* (pictured), used for centuries in China to treat back pain. The compound, dehydrocorybulbine, binds to dopamine receptors and offers longer-term relief than opiate drugs.

1864

During the American Civil War, Silas Weir Mitchell and two fellow surgeons identify an excruciating form of chronic pain that stems from damaged peripheral nerves, a condition that Mitchell calls causalgia (now called complex regional pain syndrome). Even minor injuries cause unbearable burning pain, which soldiers liken to a “red-hot file rasping the skin”. They become hypersensitive to the slightest touch; exposure to air or heat, or even the sound of a rustling newspaper increases their pain. Some are still suffering decades later.

1965

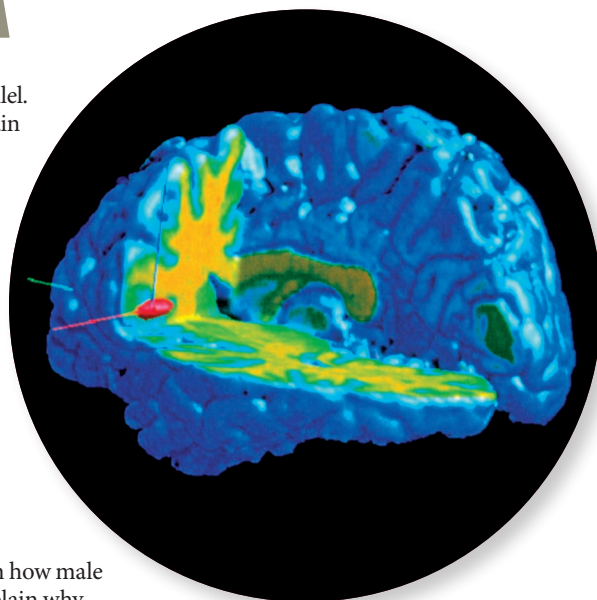
Psychologist Ronald Melzack and neuroscientist Patrick Wall propose their gate-control theory of pain. They suggest that the spinal cord has a ‘gate’ mechanism: messages from the source, other nerves and the brain converge to determine whether the gate opens to allow pain messages to reach the brain or closes to prevent them. This suggests that the perception of pain is influenced by a combination of physiological and psychological factors, such as mood. Although the details of their mechanism later prove flawed, the theory revolutionizes the field.

1973

US researchers discover a receptor in the brain through which morphine exerts its effects. This suggests that opiate drugs work by mimicking natural painkillers made by the body. Two years later, British biologists discover enkephalins, a group of endogenous opioids — or endorphins. Endorphins form part of the body's natural mechanism for managing pain, providing analgesia by reducing the perception of pain.

1991

Neuroimaging techniques reveal that pain is processed in several areas of the brain in parallel. Positron emission tomography (the brain's pain response is pictured in red) and functional magnetic resonance imaging have since provided a deeper understanding of this phenomenon, as well as how the perception of pain is influenced by emotion, experience and expectation. Some think that these techniques may make it possible to measure pain objectively and to distinguish physical from emotional pain. Imaging could help in the search for new drugs for chronic pain (see page S8).



2015

Researchers find a fundamental difference in how male and female mice process pain, helping to explain why men and women seem to feel pain differently (R. E. Sorge *et al.* *Nature Neurosci.* **18**, 1081–1083; 2015). Women are more sensitive to pain than men, are more likely to have chronic pain and respond differently to some painkillers. Past studies showed that the immune cells microglia play a key part in pain perception, but this is now found to be true only in males. T cells serve the same function in female mice (see page S7).

Seeds of neuroendocrine doubt

ARISING FROM Z. Jaunmuktane *et al.* *Nature* **525**, 247–250 (2015); doi:10.1038/nature15369

The possibility of human-to-human transmission of Alzheimer's disease has not been considered until recently. A landmark study, published in September 2015 in *Nature*, reports the probable seeding of amyloid- β (A β ; a pathological hallmark of Alzheimer's disease) from cadaveric human growth hormone (c-hGH) pituitary preparations in individuals who also developed iatrogenic Creutzfeldt–Jakob Disease (iCJD) from the same source¹. Here, we argue that c-hGH was historically used to treat rare neuroendocrine disease and that these diseases or treatment with growth hormone could result in A β pathology *per se*. Without a control group of the same patient cohort who neither received c-hGH nor developed iCJD, we believe that the authors are premature to infer an iatrogenic source of Alzheimer's disease. There is a Reply to this Brief Communication Arising by Collinge, J. *et al.* *Nature* **535**, <http://dx.doi.org/10.1038/nature18603> (2016).

Cadaveric-hGH was previously used to growth hormone deficiency (GHD) in children. In the UK, growth hormone was extracted from a cadaveric pool of approximately 400,000 pituitary glands and offered as an intramuscular treatment to selected patients. Given scarce resources and laborious extraction, its use was coordinated in the UK by a Medical Research Council working party in 1959 that was superseded by the Health Services Human Growth Hormone Committee in 1977². In the UK, 1,908 people were registered as having received c-hGH from 1959 to 1985 when it was withdrawn worldwide after an initial report documenting three deaths from CJD in the US and one in the UK. This number rose to 226 incidences worldwide in 2012³. From 1985 onwards c-hGH became obsolete after being superseded by synthetic recombinant forms.

The diverse clinical indications for c-hGH treatment until 1985 are clearly presented in a comprehensive epidemiological review which used data from the Health Services Human Growth Hormone Committee records in the UK⁴. Of the 1,908 patients registered as having received c-hGH in the UK, 1,004 (52.6%) had 'idiopathic' isolated GHD, 188 (10%) 'idiopathic' panhypopituitarism, 230 (12.1%) craniopharyngioma, 194 (10.2%) other intracranial tumour, 62 (3.3%) other neurological disease for example, hydrocephalus or post-meningitis, 62 (3.3%) genetic short stature and growth delay or both and 44 (2.3%) Turner's syndrome. The remaining causes included Prader–Willi syndrome, low birth weight, haematological disease, histiocytosis and septo-optic dysplasia. These well-recorded historical data demonstrate the heterogeneity and severity of neuroendocrine disease necessitating c-hGH treatment in the UK before 1985.

Supplementary material from the article in question reports the clinical details of all eight patients and the indication for c-hGH treatment, summarized in Table 1. All patients received c-hGH as children or adolescents and duration of treatment was 2–12 years. From an aetiological perspective, 3 of 8 subjects had panhypopituitarism (one also with mental retardation and microcephaly), 1 of 8 had craniopharyngioma and 4 of 8 are reported to have been treated for short stature alone. This is surprising as short stature alone was an uncommon indication for c-hGH before 1985, although trials were starting to emerge for this indication just before the withdrawal of c-hGH⁵. Short stature should not be confused with GHD as the former can be constitutional and the latter is a neuroendocrine disease which can lead to short stature. We therefore believe it likely that the 'short stature' group disguises a more complicated clinical picture.

At the other end of the clinical spectrum, one of the four patients reported as having moderate to severe grey matter and vascular A β pathology had a craniopharyngioma. These unusual pituitary tumours can be large, often require neurosurgery and/or cranial radiation and are prone to recurrence and multiple pituitary hormone deficiencies. Notably, the one patient in the sporadic CJD control group (aged ~65 years) who also had similar levels of cerebral amyloid angiopathy to the iCJD group was labelled as an outlier and it was stated that the subject had a surgical intervention 40 years before death (Jaunmuktane *et al.*¹ Extended Data Figs 2 and 3).

Jaunmuktane and colleagues' article was preceded by a case series from the same research group of the clinical, imaging, molecular and autopsy findings of the whole iCJD cohort ($n = 22$)⁶. One patient had a history of long-term cognitive problems and additional MRI findings included one patient with septo-optic dysplasia and partial agenesis of the corpus callosum. It is not clear whether these particular patients were the ones included in the current article, but together with the historical records and the supplementary clinical details provided, highlight our argument that these patients had complex, rare and potentially severe neuroendocrine disease.

With this in mind, an alternative hypothesis contributing to the findings is that these pre-existing and underlying conditions could by themselves lead to A β pathology and abnormal brain structure. Growth hormone is known to have various cognitive effects on the brain and isolated GHD in children can lead to lower IQ, impaired cognition, reduction brain volumes and white matter abnormalities^{7,8}. If GHD continued into adulthood (as it usually does) and was untreated then significant cognitive impairments are likely to have persisted⁹.

Table 1 | Reported indications for GH treatment and incidence of A β in study participants

Patient	Sex	History	Substantial A β	Focal A β	A β angiopathy
1	M	GH as a child for ~5 years for panhypopituitarism	No	Yes	No
2	F	GH as a child for ~10 years for panhypopituitarism, mental retardation and microcephaly	No	No	No
3	M	GH as a child for ~5 years for short stature	No	Yes	No
4	M	GH as a child for short stature from ~7 years of age for ~12 years	Yes	No	Yes
5	F	Craniopharyngioma aged ~10 years, GH from 11 years of age for ~8 years	Yes	No	No
6	M	GH for short stature from ~9 years of age for ~5 years	Yes	No	Yes
7	M	GH as an adolescent for ~3 years for panhypopituitarism	No	No	No
8	M	GH for restricted growth from 16 years of age for ~2 years	Yes	No	Yes

Summary table of the clinical indications for c-hGH treatment in childhood and incidence of A β for each patient (adapted from article text and supplementary information of ref. 1). GH, growth hormone.

We now know that traumatic brain injury (TBI) is a common cause of GHD in children¹⁰ and this may not have been widely acknowledged before 1985. TBI has been associated with widely distributed A β deposition in post-mortem brain tissue¹¹. Other neurological diseases including epilepsy, fragile X syndrome, Down's syndrome and Parkinson's disease have all been linked with significant increased levels of amyloid precursor protein (APP) and A β protein, questioning the specificity of this neuropathology to Alzheimer's disease. A β deposition may result from common underlying processes such as neuroinflammation, microglial activation and genetic factors¹².

The authors comment on the striking cerebral amyloid angiopathy in this patient group and hypothesize that these individuals "would be at increasing risk of cerebral haemorrhages had they lived longer". However, large long-term follow up studies of children who received recombinant growth hormone from 1985 onwards (that is, no risk of human-human transmission) were found to have a five- to seven-fold increased risk of cerebrovascular disease and, in particular, subarachnoid haemorrhage, suggesting that this cohort have predisposing factors leading to increased risk of intracranial haemorrhage, irrespective of having received c-hGH^{13,14}.

In conclusion, the authors present an interesting finding worthy of further investigation, but we feel that caution must be exercised in interpreting the result, especially in the absence of an appropriate control group. We believe it is also important to present a neuroendocrinological perspective and consider plausible alternative hypotheses to offer a balanced view on this high-profile communication.

Claire Feeney¹, Gregory P. Scott¹, James H. Cole¹,
Magdalena Sastre¹, Anthony P. Goldstone¹ & Robert Leech¹

¹Division of Brain Sciences, Imperial College London, Hammersmith Hospital Campus, London W12 0NN, UK.
email: c.feeney@imperial.ac.uk

Received 7 December 2015; accepted 16 April 2016.

Collinge *et al.* reply

REPLYING TO C. Feeney *et al.* *Nature* **535**, <http://dx.doi.org/10.1038/nature18602> (2016)

In the accompanying Comment¹ Feeney *et al.* argue that we have not excluded other explanations of our finding of extensive amyloid- β (A β) deposition in relatively young individuals who received extracts of human pituitary glands². They argue that this may have been a consequence of their underlying pathology that led to cadaveric pituitary-derived human growth hormone (c-hGH) treatment and propose that growth hormone deficiency and A β pathology in these patients may have resulted from traumatic brain injury (TBI).

First, we did not argue that the patients we described had developed Alzheimer's disease, rather we reported the neuropathological finding of A β pathology. Our study was observational rather than an epidemiological or experimental. These patients, as acknowledged by the authors who present a table summarizing our clinical findings, received c-hGH for various reasons and those who developed A β deposition were treated for pathogenetically unrelated conditions, including short stature of no obvious cause, making a common mechanism leading to A β pathology unlikely. The patients we reported were referred to us with CJD. We think it is unlikely that we missed significant common aspects of the medical history as all patients were under our direct care, medical records were reviewed and this issue

1. Jaunmuktane, Z. *et al.* Evidence for human transmission of amyloid- β pathology and cerebral amyloid angiopathy. *Nature* **525**, 247–250 (2015).
2. Milner, R. D. *et al.* Experience with human growth hormone in Great Britain: the report of the MRC Working Party. *Clin. Endocrinol.* **11**, 15–38 (1979).
3. Brown, P. *et al.* Iatrogenic Creutzfeldt-Jakob disease, final assessment. *Emerg. Infect. Dis.* **18**, 901–907 (2012).
4. Buchanan, C. R., Preece, M. A. & Milner, R. D. Mortality, neoplasia, and Creutzfeldt-Jakob disease in patients treated with human pituitary growth hormone in the United Kingdom. *Br. Med. J.* **302**, 824–828 (1991).
5. Buchanan, C. R., Law, C. M. & Milner, R. D. Growth hormone in short, slowly growing children and those with Turner's syndrome. *Arch. Dis. Child.* **62**, 912–916 (1987).
6. Rudge, P. *et al.* Iatrogenic CJD due to pituitary-derived growth hormone with genetically determined incubation times of up to 40 years. *Brain* **138**, 3386–3399 (2015).
7. Nyberg, F. & Hallberg, M. Growth hormone and cognitive function. *Nat. Rev. Endocrinol.* **9**, 357–365 (2013).
8. Webb, E. A. *et al.* Effect of growth hormone deficiency on brain structure, motor function and cognition. *Brain* **135**, 216–227 (2012).
9. Falletti, M. G., Maruff, P., Burman, P. & Harris, A. The effects of growth hormone (GH) deficiency and GH replacement on cognitive performance in adults: a meta-analysis of the current literature. *Psychoneuroendocrinology* **31**, 681–691 (2006).
10. Rose, S. R. & Auble, B. A. Endocrine changes after pediatric traumatic brain injury. *Pituitary* **15**, 267–275 (2012).
11. Graham, D. I., Gentleman, S. M., Lynch, A. & Roberts, G. W. Distribution of beta-amyloid protein in the brain following severe head injury. *Neuropathol. Appl. Neurobiol.* **21**, 27–34 (1995).
12. Westmark, C. J. What's hAPPening at synapses? The role of amyloid β -protein precursor and β -amyloid in neurological disorders. *Mol. Psychiatry* **18**, 425–434 (2013).
13. Carel, J. C. *et al.* Long-term mortality after recombinant growth hormone treatment for isolated growth hormone deficiency or childhood short stature: preliminary report of the French SAGhE study. *J. Clin. Endocrinol. Metab.* **97**, 416–425 (2012).
14. Poidvin, A. *et al.* Growth hormone treatment for childhood short stature and risk of stroke in early adulthood. *Neurology* **83**, 780–786 (2014).

Author Contributions C.F. wrote the article. G.P.S. contributed to writing the article. J.H.C., M.S., A.P.G. and R.L. revised the manuscript and contributed to its content.

Competing Financial Interests Declared none.

doi:10.1038/nature18602

was explored in detail with families. We identified no publications that report a causal relationship between panhypopituitarism, short stature, or craniopharyngioma and Alzheimer's disease or increased A β deposition. Multiple series of age-matched controls have been published previously and our cohort was compared to the largest series³. We did not have access to autopsies on patients treated after 1985 with synthetic growth hormone. We agree such a comparison will be of interest although difficult as these necessarily represent different age cohorts.

One alternative hypothesis proposed by the authors is that TBI in childhood may have been responsible for the A β deposition. There is no evidence that these patients had TBI in early childhood and thus it remains entirely speculative that the A β pathology may have been a long-term consequence of trauma.

Feeney *et al.* also refer to a study of 6,874 patients treated with synthetic growth hormone due to isolated idiopathic growth hormone deficiency or short stature, where 11 patients developed an intracerebral haemorrhage, ischaemic stroke or subarachnoid haemorrhage⁴. None of these patients had pathologically confirmed cerebral amyloid angiopathy (CAA) and five were known to have aneurysm, vascular

malformation or arterial dissection, an unsurprising finding in such a large cohort.

The rationale of our publication was to raise the awareness that human-to-human transmission of proteopathic seeds may not be restricted to prions. Investigating the role of seeded protein aggregation (often referred to as 'prion-like' mechanisms) is one of the most active current areas of neurodegeneration research⁵. There is already a substantial body of experimental data *in vitro* and *in vivo* demonstrating A β seeding, including evidence that peripheral inoculation of laboratory mice with Alzheimer's disease brain extracts leads to CAA. We agree with Feeney *et al.* that our study does not prove human transmission of A β pathology and further studies are needed. As we note, it will be important to examine archived batches of c-hGH for presence of A β seeds by animal inoculation studies and this work is planned. In addition, we argued in our manuscript that our findings should also prompt investigation of whether other known iatrogenic routes of prion transmission (unrelated to c-hGH treatment and/or the underlying conditions for which it was given) may also be relevant to A β seeding. The most common other cause of iatrogenic CJD is use of dura mater grafts. Importantly, an Austrian/Swiss series of dura mater graft recipients has now been reported which also found frequent vascular and parenchymal A β pathology suggesting that A β seeding can occur during procedures unrelated to neuroendocrine disease or treatment⁶.

John Collinge^{1,2,3}, Zane Jaunmuktane⁴, Simon Mead^{1,2,3}, Peter Rudge^{1,3} & Sebastian Brandner^{2,4}

¹Medical Research Council Prion Unit, Queen Square, London WC1N 3BG, UK.

²Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK.

email: s.brandner@ucl.ac.uk

³National Prion Clinic, The National Hospital for Neurology and Neurosurgery, University College London Hospitals NHS Foundation Trust, Queen Square, London WC1N 3BG, UK.

⁴Division of Neuropathology, The National Hospital for Neurology and Neurosurgery, University College London Hospitals NHS Foundation Trust, Queen Square, London WC1N 3BG, UK.

1. Feeney, C. *et al.* Seeds of neuroendocrine doubt. *Nature* **535**, <http://dx.doi.org/10.1038/nature18602> (2016).
2. Jaunmuktane, Z. *et al.* Evidence for human transmission of amyloid- β pathology and cerebral amyloid angiopathy. *Nature* **525**, 247–250 (2015).
3. Braak, H. & Braak, E. Frequency of stages of Alzheimer-related lesions in different age categories. *Neurobiol. Aging* **18**, 351–357 (1997).
4. Poidvin, A. *et al.* Growth hormone treatment for childhood short stature and risk of stroke in early adulthood. *Neurology* **83**, 780–786 (2014).
5. Jucker, M. & Walker, L. C. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* **501**, 45–51 (2013).
6. Frontzek, K., Lutz, M. I., Aguzzi, A., Kovacs, G. G. & Budka, H. Amyloid- β pathology and cerebral amyloid angiopathy are frequent in iatrogenic Creutzfeldt-Jakob disease after dural grafting. *Swiss Med. Wkly.* **146**, w14287 (2016).

doi:10.1038/nature18603

CELL DIVISION

A sticky problem for chromosomes

The role of Ki-67 in mitotic cell division has been a mystery. Extensive imaging reveals that this highly positively charged protein coats chromosomes to prevent them from coalescing. [SEE LETTER P.308](#)

CLIFFORD P. BRANGWYNNE
& JOHN F. MARKO

DNA is often imagined as an abstract string of genetic information, but in fact it is folded up by proteins into chromosomes, which are reminiscent of the dense polymeric materials that are often studied in physics and chemistry. The physical and chemical properties of this material are particularly crucial during cell division (mitosis), when the cell's chromosomes are disentangled from one another and compacted into tidy packages that segregate into two daughter cells. But after more than a century of study, the compaction and individualization of mitotic chromosomes remains poorly understood. On page 308, Cuylen *et al.*¹ report that a protein called Ki-67 coats the surface of chromosomes, providing a barrier that keeps them apart during mitosis.

During prophase, the initial stage of mitosis, replicated chromosomes are compacted into thick fibres. The membrane that surrounds the nucleus subsequently breaks down and the chromosomes separate from one another in the cytoplasm, eventually aligning along the centre of the cell during metaphase. This alignment is organized by a microtubule-based structure called the mitotic spindle, to which the chromosomes are attached. At the end of mitosis, the spindle pulls a set of chromosomes to each pole of the dividing cell.

One might expect chromosomes to coalesce rather than separate after the breakdown of the nuclear membrane, because the molecules responsible for chromosome compaction are by themselves unable to distinguish different chromosomes. Many other polymeric assemblies in cells — including RNA–protein (RNP) bodies such as P granules and nucleoli — do exhibit liquid-like coalescence on contact with each other². Cuylen *et al.* questioned what mechanisms the cell may use to keep chromosomes from sticking together when they are released into the cytoplasm.

The authors used automated live-cell imaging to analyse the effect of removing different proteins from cells. Out of more than 1,000 candidates tested, only cells lacking Ki-67 showed a severe defect in chromosome separation — their chromosomes were no longer visible as

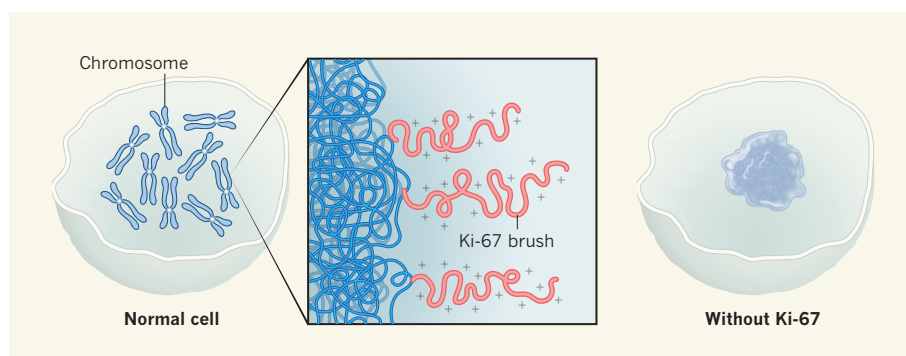


Figure 1 | Ki-67 divides the crowd. During cell division, chromosomes, which are in the form of dense cylindrical units, separate from one another. Cuylen *et al.*¹ report that chromosomes are kept apart by the protein Ki-67, which acts as a highly charged brush — one end binds to the chromosome and the positively charged protein extends into the cytoplasm, repelling other chromosomes. If Ki-67 is absent, chromosomes coalesce and cell division is delayed. (Adapted from ref. 1.)

distinct cylindrical units, but were instead stuck together in an amorphous mass. The resulting glob of chromosomes was relatively impenetrable to the growing mitotic spindle, hampering progression through mitosis.

Ki-67 is a large protein that has long been known to associate with the surface of chromosomes³. It is used as a marker of proliferation in cell biology and to assess the growth of tumour cells during cancer diagnostics. What is Ki-67 doing to keep chromosomes distinct during mitosis? Cuylen *et al.* point out that Ki-67 is a highly positively charged protein and is predicted to be mostly structurally disordered. The authors report that amino-acid substitutions or deletions in the middle or amino-terminal regions of Ki-67 had little or no effect on chromosome separation. However, the carboxy-terminal end of the protein was essential for its function.

In a set of simple but well-designed imaging experiments, Cuylen and colleagues showed that Ki-67 sits on the chromosome surface with the C-terminal end binding the chromosome and the other end sticking out into the cytoplasm, rather like bristles on a brush (Fig. 1). The authors observed a direct correlation between the amount of Ki-67 on the chromosome surface and the average spacing between chromosomes, which increased to around one micrometre when Ki-67 was at its highest levels. Thus, Ki-67 forms a brush-like arrangement on the

surface of chromosomes, apparently providing a physical and electrostatic barrier that keeps chromosomes apart.

The behaviour of Ki-67 is reminiscent of that of a class of molecules known as surfactants, which are named after their 'surface-active' behaviour. As with Ki-67, surfactant molecules have ends with different affinities. Amphiphilic phospholipids are a ubiquitous example in biology — their charged ends prefer water, whereas their hydrocarbon tails prefer oil. The tails can also stick to themselves, leading to the formation of phospholipid bilayers, such as those that make up cell membranes. More generally, surfactants assemble at interfaces, and are often used to keep droplets or particles from aggregating⁴.

Protein surfactants have arisen before in biology: for example, lipoprotein complexes keep the alveoli in lungs from sticking together and collapsing⁵. Cuylen and colleagues' work indicates that proteins also act as surfactants in the cell by stabilizing large macromolecular assemblies. Because there are many other types of membrane-less body in the cell (for example, RNP bodies and large signalling assemblies), it is possible that protein surfactants have a similar organizational role in these other structures. Between cell divisions, Ki-67 is located in nucleoli, in which the cell's protein-synthesizing machinery is assembled; perhaps it has a yet-to-be-discovered surfactant role there, helping to tune the surface

tensions of the nucleolus and thus facilitate subcompartmental organization⁶.

Cuylen and colleagues' work opens up several avenues for investigation. It remains unclear why Ki-67 comes into play only when chromosomes are in the cytoplasm and not during prophase — how are the mechanisms that underlie prophase chromosome compaction and individualization⁷ integrated with Ki-67's assembly and function? The brush-like self-assembly of Ki-67 is particularly intriguing because chromosomal DNA should be negatively charged and so attract positively charged proteins, and yet Ki-67 is oriented away from the chromosome surface.

And given the nanometre-scale distances over which electrostatic interactions occur, it will be interesting to define exactly how increasing the amount of Ki-67 increases interchromosome separation by up to a micrometre. These and other sticky chromosome questions are no doubt surfacing in many researchers' minds. ■

Clifford P. Brangwynne is in the Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, USA. **John F. Marko** is in the Departments of Molecular Biosciences and of Physics & Astronomy, Northwestern

University, Evanston, Illinois 60208, USA.
e-mails: cbrangwy@princeton.edu;
john-marko@northwestern.edu

1. Cuylen, S. *et al.* *Nature* **535**, 308–312 (2016).
2. Bergeron-Sandoval, L. P., Safaee, N. & Michnick, S. W. *Cell* **165**, 1067–1079 (2016).
3. Verheijen, R. *et al.* *J. Cell Sci.* **92**, 531–540 (1989).
4. Rosen, M. J. *Surfactants and Interfacial Phenomena* 3rd edn (Wiley, 2004).
5. Griesse, M. *Eur. Resp. J.* **13**, 1455–1476 (1999).
6. Feric, M. *et al.* *Cell* **165**, 1686–1697 (2016).
7. Goloborodko, A., Imakaev, M. V., Marko, J. F. & Mirny, L. *eLife* <http://dx.doi.org/10.7554/eLife.14864> (2016).

This article was published online on 29 June 2016.

MATERIALS SCIENCE

Magnetic nanoparticles line up

Certain bacteria contain strings of magnetic nanoparticles and therefore align with magnetic fields. Inspired by these natural structures, researchers have now fabricated synthetic one-dimensional arrays of such particles.

DAMIEN FAIVRE & MATHIEU BENNET

Simple organisms called magnetotactic bacteria passively orient along magnetic field lines like the needle of a compass¹. The orientation is dictated by the presence of a string of nanoparticles composed of magnetite (a magnetic mineral made of iron and oxygen atoms). Inspired by such bacteria, Jiang *et al.*² report in *Advanced Materials* how one-dimensional arrays of magnetite nanoparticles can be prepared synthetically, with potential applications in electronic devices and biotechnology.

Assembling magnetic nanoparticles is a challenging task, at least for humans. The main reason is that the north pole of one magnet is attracted to the south pole of another magnet. Therefore, linear configurations of magnetic nanoparticles are only metastable, because the nanoparticles at the end of the chains also attract each other. As a result, ring structures are fully stable³, but are not effective as compasses because they are isotropic (no matter how you rotate them, their magnetic properties are the same).

To stabilize several magnetic nanoparticles in a 1D configuration, magnetotactic bacteria have evolved a scaffold: one protein forms a backbone along which the particles attach, thanks to another 'anchor' protein¹. To reproduce this strategy in the laboratory, chemists have proposed using templates to align magnetic nanoparticles into 1D arrangements⁴. However, even the cleverest technical approaches have not yet led to the uniform

ordering of particles observed in magnetotactic bacteria. Impressively, though, Jiang and colleagues' template-directed technique can yield millimetre-long chains of magnetite particles (with diameters of 150 nm) that are assembled nearly perfectly edge to edge.

The authors begin by fabricating a template that consists of an array of elongated ridges (each 2 µm wide) separated by gaps (Fig. 1). The tops of the ridges are made to be hydrophilic — they have a strong affinity for water — and the base of the template is hydrophobic (water-repellent). A suspension of hydrophilic magnetic nanoparticles in water is then sandwiched between the template and a flat, inflexible target substrate. The hydrophilic top of the template attracts the suspension, whereas the base repels it, creating a thin film. The water is then evaporated away, and a mechanism known as the coffee-ring effect⁵ (named after the process in which ring-shaped deposits are produced after coffee has been spilt) causes chains of particles to be deposited along the ridges, forming 1D magnetic arrays.

Jiang and colleagues' approach combines simplicity and versatility. For example, the authors show that they can control the width of the chains by varying the concentration of the particles used, and that this tunes the properties of the magnetic array. Most importantly, the magnetic properties along the long axis of the chains differ from those perpendicular to the axis. Such directional behaviour is an essential property of 'magnetosensitive' materials and is a characteristic shared with

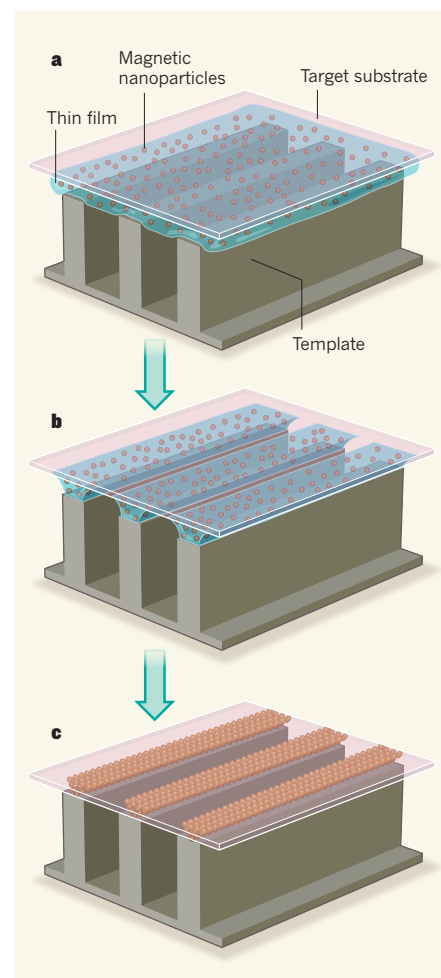


Figure 1 | The assembly of one-dimensional magnetic arrays. Jiang *et al.*² report a method for fabricating lines of magnetic nanoparticles. **a**, The authors first prepare a template consisting of elongated ridges and gaps, and a flat target substrate. A suspension of magnetic nanoparticles in water is then inserted between the template and the substrate. The suspension forms a thin film because it is attracted to the tops of the ridges, which are hydrophilic, whereas the rest of the template is hydrophobic. **b**, The water is evaporated away and the film divides into sections along the ridges of the template. **c**, Once the water has fully evaporated, a 1D arrangement of the nanoparticles forms. (Adapted from ref. 2.)

aligned magnetotactic bacteria deposited on a substrate⁶. This feature is what makes 1D magnetic structures so fascinating and potentially useful: 1D arrays have biotechnological and medical applications⁷, and possible uses in actuation devices⁸ such as nanorobots.

A potential next step will be to fabricate magneto-responsive materials that would produce a signal or a force when an external magnetic field is changed (typically, rotated). Several approaches could be investigated, such as depositing the particles on a flexible substrate, which would allow the magnetic force generated when an array is placed in a magnetic field to be converted into mechanical work (for example, contraction or expansion of the substrate). Restriction of the magnetic array to a single line would also be ideal for the engineering of devices.

Another potential development concerns miniaturization. The particles synthesized by bacteria are optimized because, although only around 30 nm in diameter, they are large

enough for their magnetic properties to be cumulative. The particles currently used by the authors are 150 nm in diameter, but are aggregates of smaller particles that do not exhibit cumulative magnetic properties. Therefore, the same magnetic properties could be achieved with particles as small as 5–10 nm. Using particles that are the same size as those in magnetotactic bacteria would change the magnetic properties of the arrays, enhancing their sensitivity and enabling rotation of the chains by low magnetic fields (potentially, of similar strengths to that of Earth's magnetic field).

Magnetite nanoparticles are special because their magnetic properties depend in part on their shape. Therefore, using nanocubes instead of round particles would again lead to modified magnetosensitive properties. The fabrication of 2D or 3D patterns will surely also have many applications. Crucially, with all of these ideas in mind, the method developed by Jiang and collaborators has the potential to aid

the development of functional nanomaterials that mimic the naturally occurring perception of magnetic fields. ■

Damien Faivre and Mathieu Bennet are at the Max Planck Institute of Colloids and Interfaces, Potsdam-Golm Science Park, 14476 Potsdam, Germany.
e-mails: damien.faivre@mpikg.mpg.de; mathieu.bennet@mpikg.mpg.de

1. Faivre, D. & Utkar Godec, T. *Angew. Chem. Int. Edn* **54**, 4728–4747 (2015).
2. Jiang, X. *et al. Adv. Mater.* <http://dx.doi.org/10.1002/adma.201601609> (2016).
3. Kiani, B., Faivre, D. & Klumpp, S. *New J. Phys.* **17**, 043007 (2015).
4. Yuan, J., Xu, Y. & Müller, A. H. E. *Chem. Soc. Rev.* **40**, 640–655 (2011).
5. Deegan, R. D. *et al. Nature* **389**, 827–829 (1997).
6. Alphandéry, E. *et al. ACS Nano* **3**, 1539–1547 (2009).
7. Ghosh, D. *et al. Nature Nanotechnol.* **7**, 677–682 (2012).
8. Mishra, S. R., Dickey, M. D., Velebv, O. D. & Tracy, J. B. *Nanoscale* **8**, 1309–1313 (2016).

scenarios, Thackeray *et al.* forecast that, by 2050, primary consumers will have shifted the timing of their phenological events twice as much as will species at other trophic levels.

But why do secondary consumer species respond more weakly to temperature compared with primary consumers or producers? After all, both predators and prey will use temperature as a cue to time their phenology.

The reason is that species at different trophic levels do not rely on exactly the same temperature cue. Different species respond to temperature at different times of year. Take, for instance, a winter moth egg of 2 millimetres in diameter, which has been lying dormant on a 40-metre-tall oak tree for months. It manages to hatch within days of its host's bud burst. Obviously, these two species will have radically different physiological mechanisms underlying their phenological events and will use different cues to determine the timing of these events. So, although the two sets of cues will be correlated, the cue used by the consumer (the moth in this case) will always be, to some extent, unreliable. Theoretical work³ shows that this imperfect cue reliability means consumers will evolve a less temperature-sensitive phenology than will the species at the trophic level they rely on.

As for understanding the consequences of the phenological mismatches caused by climate change, it is notable that almost all previous research has focused on simple two-species interactions of a predator and its prey, or a herbivore and its host plant. But despite the clear effects of the phenological mismatches on population numbers, such effects can be buffered by ecological mechanisms such as density dependence⁴ — whereby the success rate of an individual increases when the number of individuals declines.

PHENOLOGY

Interactions of climate change and species

At what times of year are phenological events across species sensitive to climatic variables, and how sensitive are they? Answers to these questions emerge from the analysis of a wealth of long-term data sets. [SEE ARTICLE P.241](#)

MARCEL E. VISSER

Spring is an exciting time of year, with flowers coming into bloom, bees emerging from their winter rest and migrant birds returning to their breeding grounds. In the United Kingdom, all such phenological events have been monitored for decades — across the seasons — by professionals and citizen scientists alike (Fig. 1). And although many of these schemes started well before we recognized that the world is warming, these long time series now provide a valuable infrastructure to document the impact of global climate change on nature. On page 241, Thackeray *et al.*¹ use no fewer than 10,003 phenological data sets, all collected over periods of at least 20 years, to quantify the unequal rates at which phenology has shifted in different species groups. Climate change thus leads to disruption of the phenological match between species, and so to ecological relationship problems.

Climate change leads to unequal shifts in phenology across species². To understand why, Thackeray and colleagues studied the sensitivity of phenological events to two climate variables in a wide range of species. Specifically,

for each of their 10,003 phenological time series, they asked at what times of year species are sensitive to changes in temperature and precipitation, and how sensitive events in their phenological cycles are to these variables. Both questions are relevant, because climate has changed — and will continue to change — in a non-uniform manner over the course of the year. Some periods warm faster than others, and two species with equal temperature sensitivities but at different times may thus shift their phenological events at different rates.

On the whole, the authors found that species at different trophic levels (positions) of the food chain did not differ in the times of year at which they were sensitive to annual variations in temperature, but did vary in how sensitive they were. The phenology of species high up the food chain (the secondary consumers) is much less sensitive to temperature than is that of species at the base of the chain — the primary producers and primary consumers, which are twice as sensitive to temperature. Secondary consumers were also less sensitive to variations in precipitation. When combining the species sensitivities with climate



Figure 1 | Citizen scientist. Thackeray and colleagues' study¹ of how phenology has shifted for different species groups in response to climate change is based on previous records of events collected by citizen scientists, as well as professionals.

A more overarching, but also more challenging, approach would be to turn away from these two-species interactions and to look at the effects on the entire food web⁵. How are the strengths of the links in a food web affected by phenological mismatches? What happens if the phenology of species at one trophic level shifts more than that of species at another? Does this lead to the loss of some links and the formation of others? Does this

destabilize the web? Such analyses would be a stepping stone from studying the phenological shifts of species to understanding the effects of climate change on ecosystem function⁶.

But, to complicate matters further, species' climate sensitivity is not fixed. The phenological mismatches lead to selection on the timing of phenological events. And, because phenology is often heritable, this leads to genetic change in sensitivity⁷.

It will be a major challenge to combine genetic change with a food-web approach, and to include the necessary detailed climatological projections. But it is one that must be undertaken to forecast the effects of climate change, through phenological responses, on ecosystem function. What is clear is that long time series, such as the 10,003 analysed in the present paper, are essential for this. Therefore, professional and citizen scientists — who together made the 379,000 individual phenological observations on which these time series are based — need to be encouraged and facilitated to keep up their good work. The additional advantage is that observing phenological shifts in, sometimes literally, your own backyard drives the message of global climate change home. ■

Marcel E. Visser is in the Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen 6700, the Netherlands.

e-mail: m.visser@nioo.knaw.nl

1. Thackeray, S. J. *et al. Nature* **535**, 241–245 (2016).
2. Thackeray, S. J. *et al. Glob. Change Biol.* **16**, 3304–3313 (2010).
3. Gienapp, P., Reed, T. E. & Visser, M. E. *Proc. R. Soc. B* **281**, 20141611 (2014).
4. Reed, T. E., Grøtan, V., Jenouvrier, S., Sæther, B.-E. & Visser, M. E. *Science* **340**, 488–491 (2013).
5. Stevenson, T. J. *et al. Proc. R. Soc. B* **282**, 20151453 (2015).
6. Memmott, J., Craze, P. G., Waser, N. M. & Price, M. V. *Ecol. Lett.* **10**, 710–717 (2007).
7. Visser, M. E. *Proc. R. Soc. B* **275**, 649–659 (2008).

This article was published online on 29 June 2016.

ASTROPHYSICS

Variable snow lines affect planet formation

Observations of the disk of dust and gas around a nascent star reveal that the distance from the star at which water in the disk forms ice is variable. This variation might hinder the formation of planets. [SEE LETTER P.258](#)

BRENDA MATTHEWS

The behaviour of ices in circumstellar disks — the rings of dust and gas that surround young stars — is thought to be crucial to the amalgamation of solid bodies up to the sizes of planets in star systems. The radial distance from a star at which a volatile substance, such as water, sublimates within a disk is called its ice or snow line¹. On page 258, Cieza *et al.*² report the detection of the water snow line in the disk around a forming star, V883 Orionis (V883 Ori). Their findings

reveal that the radial locations of snow lines can vary dramatically with the rate at which the star accretes matter. This has major implications for the processes that assemble material into planets.

The water snow line is the most elusive of all snow lines, because the sublimation temperature of water is high (150 kelvin for pressures at typical snow-line distances around solar-type stars) compared with those of other volatile compounds such as carbon monoxide (20 K) or methane (65 K), placing it comparatively close to the star. But V883 Ori belongs to the

FU Orionis class of stars³, whose accretion rates vary over time. Accretion of material onto the star can occur in relatively short (up to 100 years) bursts of high activity, followed by periods of diminished accretion and relative quiescence^{4,5}. In its enhanced accretion state, the luminosity of V883 Ori increases to about 60 times that expected on the basis of the star's mass and estimated age³, resulting in increased temperatures across its circumstellar disk. Cieza and colleagues were able to detect the water snow line in their study because they targeted V883 Ori during a period of particularly high accretion that made the star hotter than normal, and therefore moved the line much farther from the star.

The authors used the Atacama Large Millimeter/submillimeter Array (ALMA, the world's largest ground-based observatory) in northern Chile to observe emissions from millimetre-sized dust grains and from CO gas molecules in the disk around V883 Ori. ALMA provides unprecedented resolution and sensitivity for emissions in the microwave part of the electromagnetic spectrum. The astute reader will note that the authors did not observe water itself. Making observations of

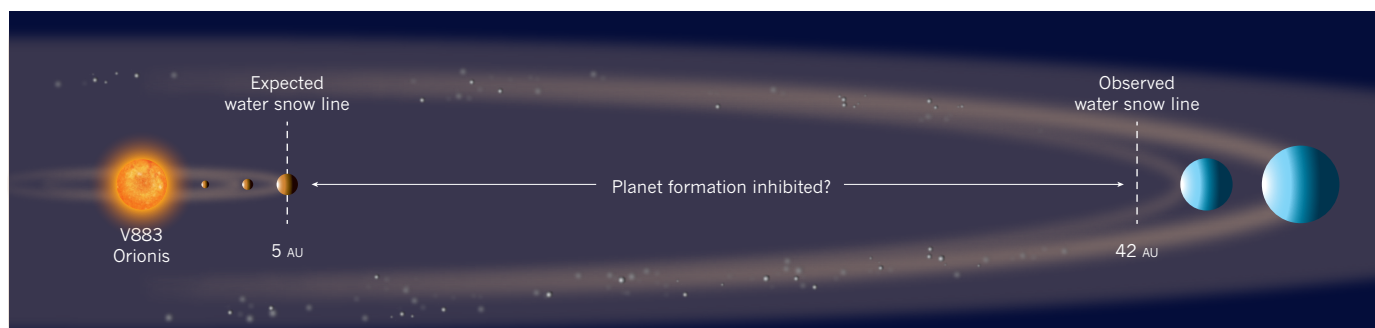


Figure 1 | Variable snow lines in the V883 Orionis system. Disks of dust and gas form around nascent stars, and the radial distance from the star at which water sublimates within a disk is called the water snow line. Cieza *et al.*² observe that the water snow line for the star V883 Orionis is located at a radial distance of 42 AU (1 AU is the distance from Earth to the Sun), whereas the expected distance, based on the star's mass and age, is about 5 AU. This

suggests that water snow lines are variable and depend on the rate at which the star is accreting matter from the disk. Because the agglomeration of icy grains is a key process in the formation of giant planets (blue), such variability might prohibit efficient giant-planet formation within the disk region bounded by the inner and outer snow lines. Rocky planets (brown) could still form within the snow line.

water from the ground is highly challenging because of the high water content of Earth's atmosphere. Direct detection of water is therefore typically the purview of space-based observatories and has been used, for example, to study the disk around TW Hydra⁶, a more-evolved star than V883 Ori.

Cieza and colleagues' image of the dust emission in the disk around V883 Ori achieves a resolution of 12 AU across the disk (1 AU is the distance from Earth to the Sun), which is impressive, given the object's distance of 414 parsecs (1,350 light years). The disk shows two distinct regions: a bright inner disk that extends to 42 AU, and a fainter outer disk extending to 125 AU. The authors show that there is a sharp transition between these two regions. Because the ALMA data are extremely sensitive and their signal-to-noise ratio is high, the authors were able to measure the spectral index — the rate of change in emission strength as a function of wavelength — and find that its value changes abruptly across the radial boundary between the inner and outer disk.

A paper⁷ last year used theoretical methods to predict such a sharp change in spectral index across the water snow line. The models that were used in that work assumed that dust grains outside the line are covered in water ice, increasing the likelihood of colliding grains sticking to each other to form centimetre-sized aggregates, whereas dust grains within the line lack an ice coating and so aggregate much less efficiently. The grains within the line therefore collide and fragment, making millimetre-sized grains more abundant.

Cieza *et al.* conclude that the transition to aggregation behaviour occurs where the temperature has dropped to below 105 K ($\pm 10\%$), consistent with the sublimation temperature of water expected in the interstellar medium. However, the authors' temperature estimate relies on an extrapolation based on the density of dust at the surface of the disk's inner region, and may be an underestimate. Future measurements of the inner disk at longer

wavelengths will improve the measured surface density, because the dust is 'optically thin' at these wavelengths, which means that emission from all the grains can be detected. Better measurements of surface density should allow improved extrapolation of the pressure, and hence of the temperature, at the radius of the snow line.

The fact that the location of the snow line can evolve with time has strong implications for planet formation. The timescales for FU Orionis-like activity outbursts are decades to centuries, much shorter than the timescales of planet-formation processes. This means that cyclical loss of icy mantles from grains in disk regions between the innermost and outermost boundaries of water snow lines might prohibit any long-term planetary growth in that region (Fig. 1). This in turn would confound

models that predict the slow formation of rocky planets within the snow line, and rapid gas-giant formation outside it. ■

Brenda Matthews is with the Herzberg Astronomy and Astrophysics Programs, National Research Council of Canada, Victoria V9E 2E7, Canada.
e-mail: brenda.matthews@nrc-cnrc.gc.ca

1. Hayashi, C. *Prog. Theor. Phys. Suppl.* **70**, 35–53 (1981).
2. Cieza, L. A. *et al. Nature* **535**, 258–261 (2016).
3. Strom, K. & Strom, S. *Astrophys. J.* **412**, L63–L66 (1993).
4. Herbig, G. H. *Astrophys. J.* **217**, 693–715 (1977).
5. Hartmann, L. & Kenyon, S. J. *Annu. Rev. Astron. Astrophys.* **34**, 207–240 (1996).
6. Zhang, K., Pontopiddan, K. M., Salyk, C. & Blake, G. A. *Astrophys. J.* **766**, 82 (2013).
7. Banzatti, A. *et al. Astrophys. J.* **815**, L15 (2015).

METROLOGY

Schrödinger's cat beats a quantum limit

Quantum effects have been used in devices that measure various quantities, but not to measure electric fields. The sensitivity of an electrometer has now been boosted using the phenomenon of quantum superposition. SEE LETTER P.262

CHARLES S. ADAMS

The grand challenge of quantum technology is to engineer quantum systems to outperform their classical counterparts in every department — whether in computing, communications or sensing. On page 262, Facon *et al.*¹ take a stride towards this goal by demonstrating a quantum-engineered sensor, based on exotic atoms, that measures weak microwave-frequency electric fields with a sensitivity comparable to that of anything

achieved previously. Even more impressively, the authors do this using a quantum trick that effectively generates 'gears' to increase the sensitivity. The trick is based on a variant of the quantum superposition phenomenon represented by Schrödinger's cat in the famous thought experiment.

Independent atoms make great sensors because the motion of their outer electrons is highly sensitive to the environment, and changes in that motion can be read out using lasers. Not only are atomic sensors sensitive

and accurate, but, unlike other sensors, they do not need to be calibrated against a standard because all atoms of a given type are identical. Atomic sensing of time provides the basis for the Global Positioning System, and sensing of magnetic fields is beginning to have an impact on medicine, archaeology and security. Moreover, atomic sensing of acceleration and gravity is more accurate than any other method². It has long been known that atomic sensitivity can be further enhanced using quantum physics^{3,4} — for example, a quantum-logic atomic clock³ reported in 2010 remains one of the most accurate timekeeping devices.

But electrometry — the measurement of electric fields — has been largely missing from the repertoire of atomic sensors. Different types of electric field, such as those associated with direct or alternating currents, present different challenges for measurement. Measuring constant (direct-current) electric fields is always tricky because the motion of free charges, such as electrons in metals or charges on insulating surfaces in the system through which the field is passing, are hard to remove completely and tend to screen the field.

This is less of a problem for high-frequency alternating-current fields, such as those in the microwave or terahertz region of the electromagnetic spectrum, because the free charges cannot respond to the oscillating field (resonate) quickly enough to screen it. However, atoms tend not to respond to such fields either. Unless, that is, we use an exotic type of atom known as a Rydberg atom, in which the outer electron is excited using a laser such that it is only weakly bound and, on average, spends most of its time far away from the nucleus. Rydberg atoms are highly sensitive to electric fields, especially microwave fields that resonate with a transition between two of the atoms' excited electronic states. In the past few years, this sensitivity has been combined with techniques that allow an optical signal to be produced in response to an electric field, to measure both non-resonant⁵ and resonant fields. But until now, all electrometry has depended on essentially classical physical effects.

A remaining challenge was therefore to engineer quantum states to increase the sensitivity of an electrometer. This is addressed in Facon and colleagues' study. Not only do the authors demonstrate quantum engineering, but they also match the sensitivity record for electrometry⁶. Perhaps most surprisingly, the extreme sensitivity of their system is achieved by measuring one atom at a time, whereas previous electrometers⁶ relied on measuring many atoms to achieve a comparable performance. The most fascinating part of their work is their use of quantum superposition — a combination of two states that have widely different energy values — to obtain this result.

To understand how quantum superposition can help, consider a sensor in which the number of atoms, N , is increased — the sensitivity

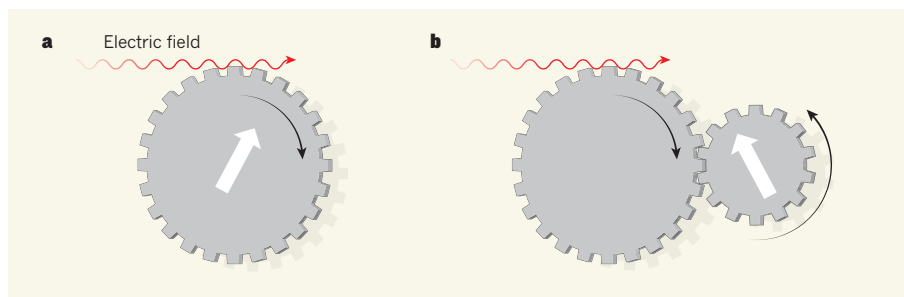


Figure 1 | Quantum effects increase sensitivity in an atomic electrometer. **a**, Electric fields can drive a transition between an electronic ground state and an excited state in atoms. Here, the sloped arrow represents a quantum superposition of excited and ground states for an atom. In a classical atomic electrometer, the electric field can be thought of as driving a wheel that rotates the arrow. The read-out of field strength depends on whether the wheel has rotated to a position at which the arrow is more up than down. The more the wheel rotates for a given electric field, the higher the sensitivity of the device. **b**, Facon *et al.*¹ have used quantum superposition to couple the first wheel to a second one, analogous to mechanical gears. Because the second wheel rotates farther than the first one for a given electric field, the sensitivity of the device is higher than that of the classical electrometer.

goes up because more signal is obtained. But the sensitivity does not rise linearly with N , because the noise that accompanies the signal also increases, with the square root of N . This is because the read-out of an atomic sensor is binary: measurements are made by counting the number of atoms in one of two states, either the ground state or an excited state. Just like counting the number of times we get heads when a coin is tossed N times, the number of events that occur within a given time frame varies, exhibiting a Poisson distribution. These counting statistics limit the precision to what is called the standard quantum limit.

The quantum trick to beat this Poissonian noise is to correlate all the atoms in a special way, for example in a quantum superposition known as a Schrödinger cat state. Such a superposition increases the energy gap between the electronic states involved in sensing, thus making the system more sensitive to external perturbation. The enhanced sensitivity is like a gearing system (Fig. 1). In this analogy, the electric field rotates a gear wheel that is coupled to another wheel through quantum superposition; the second wheel rotates faster than the gear wheel, thereby giving a more sensitive read-out. The resulting sensitivity can approach the Heisenberg limit, the maximum sensitivity that can be achieved.

Facon and colleagues use this approach in their electrometer, but rather than using N atoms, they simply use N states in one Rydberg atom. This works because the difference in energy between each adjacent state is the same, like the distance between the rungs of a ladder. The energy difference between the bottom and the N th rung is therefore the same as the energy difference between having N atoms in the ground state and N atoms on the first rung. The authors used lasers and pulsed electric fields to prepare a Schrödinger cat state that yields an N -fold energy enhancement, and hence the desired N -fold sensitivity improvement for electrometry. Preparing the cat state is the hard part, but the authors are past masters

at this, and show that they can easily beat the standard quantum limit for their sensor.

By applying their expertise in quantum felines to electrometry, Facon *et al.* open a new chapter in quantum metrology. Their sensor apparatus is currently rather bulky (metres across) and works only at particular microwave frequencies, but their results impressively demonstrate how cat states enable more-accurate measurements to be made. In addition, their work clearly puts quantum-engineered atomic electrometers right up there with the best atomic clocks and magnetometers. Further work will probably yield devices that exceed the performance of the current best electrometers, and might find applications in quantum radar and astronomy. ■

Charles S. Adams is in the Department of Physics, Durham University, Durham DH1 3LE, UK.
e-mail: c.s.adams@durham.ac.uk

1. Facon, A. *et al.* *Nature* **535**, 262–265 (2016).
2. Dickerson, S. M., Hogan, J. M., Sugarbaker, A., Johnson, D. M. S. & Kasevich, M. A. *Phys. Rev. Lett.* **111**, 083001 (2013).
3. Chou, C. W., Hume, D. B., Koelemeij, J. C. J., Wineland, D. J. & Rosenband, T. *Phys. Rev. Lett.* **104**, 070802 (2010).
4. Muessel, W., Strobel, H., Linnemann, D., Hume, D. B. & Oberthaler, M. K. *Phys. Rev. Lett.* **113**, 103004 (2014).
5. Mohapatra, A. K., Bason, M. G., Butscher, B., Weatherill, K. J. & Adams, C. S. *Nature Phys.* **4**, 890–894 (2008).
6. Sedlacek, J. A., Schwettmann, A., Kübler, H., Löw, R., Pfau, T. & Shaffer, J. P. *Nature Phys.* **8**, 819–824 (2012).

CORRECTION

In the News & Views article 'Earth science: An extended yardstick for climate variability' by Nele Meckler (*Nature* **534**, 626–628; 2016), the stalagmites shown in Figure 1 were, in fact, stalactites — the image was upside down. The image in the online article has now been replaced.

Phenological sensitivity to climate across taxa and trophic levels

Stephen J. Thackeray¹, Peter A. Henrys¹, Deborah Hemming², James R. Bell³, Marc S. Botham⁴, Sarah Burthe⁵, Pierre Helaouet⁶, David G. Johns⁶, Ian D. Jones¹, David I. Leech⁷, Eleanor B. Mackay¹, Dario Massimino⁷, Sian Atkinson⁸, Philip J. Bacon⁹, Tom M. Brereton¹⁰, Laurence Carvalho⁵, Tim H. Clutton-Brock¹¹, Callan Duck¹², Martin Edwards⁶, J. Malcolm Elliott¹³, Stephen J. G. Hall¹⁴, Richard Harrington³, James W. Pearce-Higgins⁷, Toke T. Høye¹⁵, Loeske E. B. Kruuk^{16,17}, Josephine M. Pemberton¹⁶, Tim H. Sparks^{18,19}, Paul M. Thompson²⁰, Ian White²¹, Ian J. Winfield¹ & Sarah Wanless⁵

Differences in phenological responses to climate change among species can desynchronise ecological interactions and thereby threaten ecosystem function. To assess these threats, we must quantify the relative impact of climate change on species at different trophic levels. Here, we apply a Climate Sensitivity Profile approach to 10,003 terrestrial and aquatic phenological data sets, spatially matched to temperature and precipitation data, to quantify variation in climate sensitivity. The direction, magnitude and timing of climate sensitivity varied markedly among organisms within taxonomic and trophic groups. Despite this variability, we detected systematic variation in the direction and magnitude of phenological climate sensitivity. Secondary consumers showed consistently lower climate sensitivity than other groups. We used mid-century climate change projections to estimate that the timing of phenological events could change more for primary consumers than for species in other trophic levels (6.2 versus 2.5–2.9 days earlier on average), with substantial taxonomic variation (1.1–14.8 days earlier on average).

Numerous long-term ecological changes have been attributed to climate change¹. Shifts in the seasonal timing of recurring biological events such as reproduction and migration (that is, phenological changes) have been especially well documented^{2,3}. Long-term ecosystem studies^{4–7} and global meta-analyses^{2,3,8} have demonstrated that many spring and summer phenological events now occur earlier in the year than they did previously. Substantial among-species variation in responses has fuelled concerns that key seasonal interactions among species may desynchronize over time, with potentially severe consequences for wild populations and, hence, for ecosystem functioning⁹.

Identifying systematic taxonomic and trait-based differences in phenological climate sensitivity (that is, change in seasonal timing per unit change in climatic conditions) would have substantial socio-ecological implications. Such knowledge would afford some predictability to future ecological outcomes and would identify species that represent effective sentinels of climate impact, facilitating the development of indicators and estimates of vulnerability for conservation and national adaptation programmes^{10–12}. Unfortunately, such generalizations are currently elusive.

Analytical approach and data sets

Among-species differences in phenological change may arise from two aspects of climate sensitivity. First, variation may reflect differences in physiological and behavioural responses, microclimate use, and the importance of non-climate-related cues, such as photoperiod¹³

or resource availability¹⁴. Thus, even if species are most sensitive to climate change during the same seasonal period (window), they show different phenological responses to a given climatic change. Second, co-occurring species may vary in their seasonal periods of climate sensitivity, each being typified by different levels of directional climate change^{15–17}. We conceptualize these two aspects of phenological responses as species-specific (or population-specific) climate sensitivity profiles (CSPs; Fig. 1). The CSP approach differs fundamentally from attempts to identify single ‘critical’ seasonal periods within which climatic change most strongly affects seasonal events¹⁷, by quantifying the full range of phenological responses to seasonal climatic change. We ask, “How sensitive are phenological events to temperature and precipitation change at different times of year?”. By applying this approach to a large, taxonomically diverse, national-scale data set, we discern coherent patterns within a multitude of idiosyncratic biological climate responses. We assess whether systematic differences in climate sensitivity underpin differences in phenological change among taxonomic and trophic groups in the UK⁸.

We elected against using published climate responses that may be biased in favour of species showing an effect. Instead, we analysed 10,003 long-term (≥ 20 -year) phenological time series for 812 marine, freshwater and terrestrial taxa over the period 1960–2012. Our data set aggregates many of the UK’s foremost long-term biological monitoring schemes (Supplementary Table 1), including phenological information on amphibians (spawning), birds (egg laying, migration), planktonic

¹Centre for Ecology & Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster, Lancashire LA1 4AP, UK. ²Met Office, FitzRoy Road, Exeter, Devon EX1 3PB, UK. ³Rothamsted Research, West Common, Harpenden, Hertfordshire AL5 2JQ, UK. ⁴Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB, UK. ⁵Centre for Ecology & Hydrology, Bush Estate, Penicuik, Midlothian EH26 0QB, UK. ⁶The Sir Alister Hardy Foundation for Ocean Science, The Laboratory, Citadel Hill, Plymouth, Devon PL1 2PB, UK. ⁷British Trust for Ornithology, The Nunnery, Thetford, Norfolk IP24 2PU, UK. ⁸The Woodland Trust, Kempton Way, Grantham, Lincolnshire NG31 6LL, UK. ⁹Futtie Park, Banchory, Aberdeen AB31 4RX, UK. ¹⁰Butterfly Conservation, Manor Yard, East Lulworth, Wareham, Dorset BH20 5QR, UK. ¹¹Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK. ¹²Sea Mammal Research Unit, Scottish Oceans Institute, East Sands, University of St Andrews, St Andrews, Fife KY16 8LB, UK. ¹³The Freshwater Biological Association, The Ferry Landing, Far Sawrey, Ambleside, Cumbria LA22 0LP, UK. ¹⁴University of Lincoln, Riseholme Hall, Riseholme Park, Lincoln, Lincolnshire LN2 2LG, UK. ¹⁵Aarhus Institute of Advanced Studies, Department of Bioscience and Arctic Research Centre, Aarhus University, Høegh-Guldbergs Gade 6B, DK-8000 Aarhus C, Denmark. ¹⁶Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3FL, UK. ¹⁷Research School of Biology, The Australian National University, ACT 2612 Australia. ¹⁸Faculty of Engineering and Computing, Coventry University, Priory Street, Coventry CV1 5FB, UK. ¹⁹Institute of Zoology, Poznań University of Life Sciences, Wojska Polskiego 71C, 60-625 Poznań, Poland. ²⁰University of Aberdeen, Lighthouse Field Station, George Street, Cromarty, Ross-shire IV11 8YJ, UK. ²¹People’s Trust for Endangered Species, 15 Cloisters House, 8 Battersea Park Road, London SW8 4BG, UK.

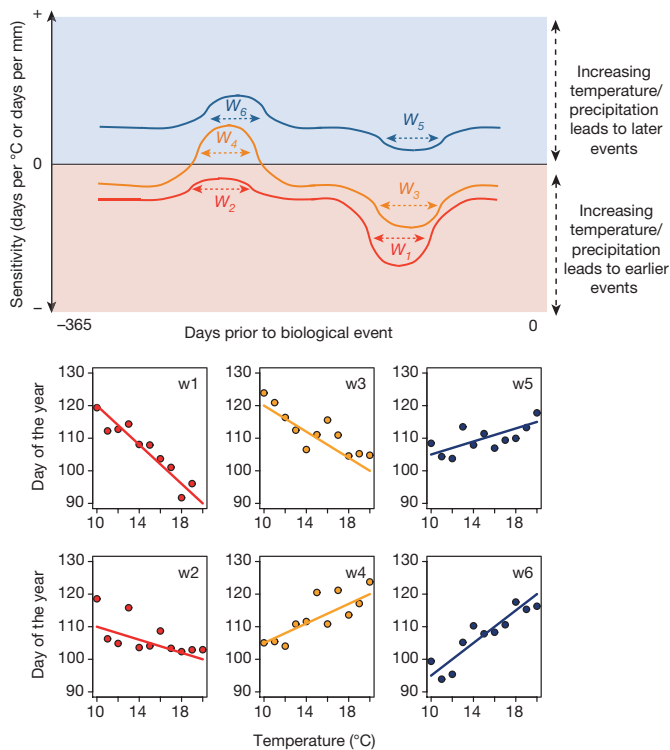


Figure 1 | Climate sensitivity profiles. Climate sensitivity is the change in seasonal timing per unit change in temperature (days per °C) or precipitation (days per mm). Irrespective of the date, increasing temperature or precipitation might always correlate with earlier (red curve; CSP type I) or later (blue curve; CSP type III) biological events, but the sensitivity to climate variation (correlation magnitude) can differ (compare w1 and w2, w5 and w6). By contrast, opposing climate–phenology correlations can occur, depending on the date at which the climate changes (orange curve, w3 and w4; CSP type II). Panels show hypothetical relationships for seasonal windows w1–w6.

crustaceans (population peaks), fish (spawning, migration), insects (flight periods), mammals (birth dates), phytoplankton (population peaks) and plants (flowering, fruiting, leafing). These taxa represent three broad trophic levels: primary producers (phytoplankton and plants), primary consumers (granivorous birds, herbivorous insects, mammals and planktonic crustaceans) and secondary consumers (predatory amphibians, birds, fish, insects, mammals and planktonic crustaceans). We spatially matched all 10,003 phenological time series with local temperature and precipitation data from a 5 × 5-km resolution gridded data set, before statistically modelling the relationship between seasonal timing and climatic variables. Between 1960 and 2012, mean UK air temperature increased in all months, and mean precipitation increased in most months (Fig. 2a).

Spatial variability in climatic change (Fig. 2b, c) necessitates local matching of phenological and climatic data sets rather than the use of regionally averaged climate data (for example, Central England Temperatures) or large-scale climatic indicators (for example, North Atlantic Oscillation). We did not make the restrictive assumption that biological events would be related to annual mean climatic conditions, or to conditions within periods based upon calendar months. Our CSP approach identified seasonal periods within which climatic change had the most positive and negative correlations with phenology (hereafter referred to as upper and lower limits of climate sensitivity, respectively). We could identify, for each phenological series, up to two seasonal periods within which climatic variation had a marked correlation with seasonal timing. The method was flexible enough to allow situations in which climatic variation within only a single period had a marked correlation with seasonal timing, and to identify seasonal windows ranging

from a few days to a whole year in length. Our analysis captured the idiosyncrasies of phenological responses, allowed us to categorize them into generic types of climate response, and is consistent with current biological understanding of climate–phenology relationships^{15,16}.

Climate response types in the UK

CSPs fall into three categories. The qualitative type of climate–phenology correlation (positive or negative) may remain consistent, irrespective of when in the year climatic change occurs. In this case only the magnitude of the phenological response differs with the time of year at which climatic variables change. The climate–phenology correlation may be consistently negative (CSP type I; Fig. 1, red curve) or positive (CSP type III; Fig. 1, blue curve). Alternatively, opposing correlations between seasonal climatic change and the timing of biological events may exist; that is, the direction and magnitude of the phenological response may vary (CSP type II; Fig. 1, orange curve). We determined CSPs for responses to temperature (CSP_{temp}) and precipitation (CSP_{precip}).

Among responses to temperature changes, CSP type II was most common (Extended Data Table 1; 69.7% of phenological series). Seasonal events were advanced by (that is, negatively correlated with) warming during one period of the year, and delayed by (that is, positively correlated with) warming in another period. After multiple testing correction, 44.8% of the observed phenological advances (but only 1.0% of delays) with warming were statistically significant ($P < 0.05$). CSP type I was the next most common response type: warming in different seasonal windows was consistently correlated with earlier seasonal events (that is, negative correlations; 24.7% of series). In this case the lower and upper limits of CSPs represent the ‘strongest’ and ‘weakest’ phenological advances with warming, respectively, and 58.1% of the ‘strongest’ responses were statistically significant ($P < 0.05$, correcting for multiple testing).

Phenological events most commonly demonstrated opposing (Fig. 1; CSP type II, 53.0% of series) or consistently positive (Fig. 1; CSP type III, 28.0% of phenological series) correlations with increasing seasonal precipitation. Although delayed phenological events may commonly be associated with higher precipitation (81.0% of events show this type of response), few of these associations were statistically significant (Extended Data Table 1).

Climate sensitivity at the UK-wide scale

We matched each phenological series with four climate variables: mean temperature during the seasonal windows at the upper and lower limits of CSP_{temp}, and similarly averaged precipitation data for the seasonal windows at the upper and lower limits of CSP_{precip}. We then combined all 10,003 phenological series and their matched climate data, and modelled the relationships between seasonal timing and climate variables using linear mixed effects (LME) models. Initially we fitted a ‘global’ model to quantify the upper and lower limits of temperature and precipitation sensitivity, averaged across all phenological events. Marine plankton data were excluded at this stage, owing to a lack of precipitation data.

Most phenological events occurred earlier with seasonal warming (average rate -2.6 days per °C; Fig. 3a and Extended Data Table 2). Variation in the strength of this correlation was similar among sites and species (random-effects variances in site and species level seasonal timing–temperature slopes were 2.1 and 1.9, respectively). Some phenological events occurred later with seasonal warming (Fig. 3a) although, in other cases, the upper limit of CSP_{temp} was in fact a ‘weak’ advance with warming. The upper limit of temperature sensitivity was more variable among species than among sites (random effects variances in species and site level seasonal timing–temperature slopes were 2.3 and 0.4, respectively). Averaged across species and populations, temperature responses were most consistent with CSP type II.

Most phenological events showed opposing responses to increasing seasonal precipitation (Fig. 1; CSP type II). The tendency for

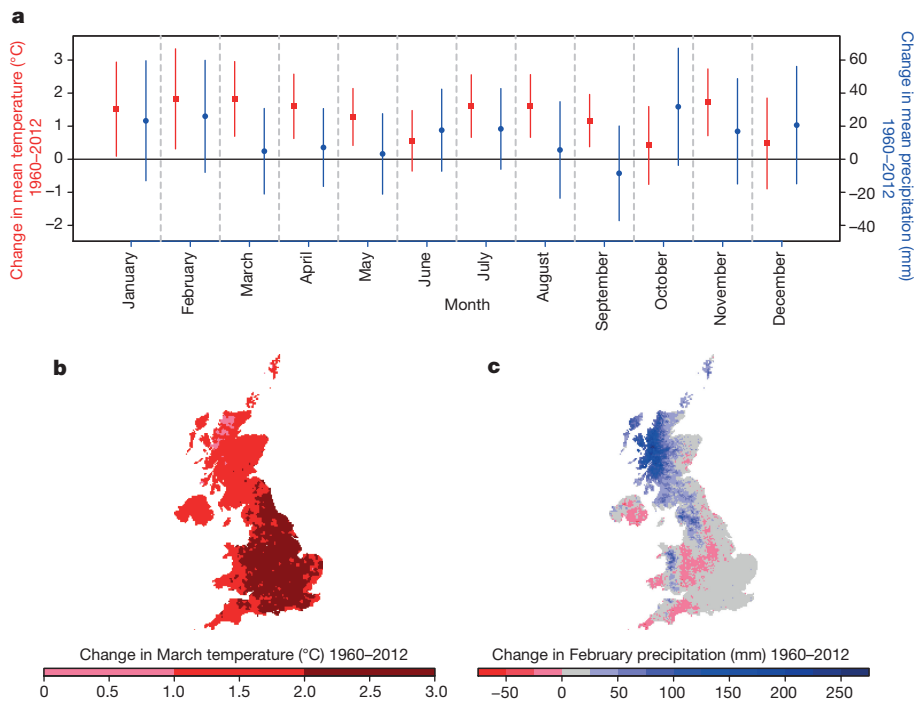


Figure 2 | Climatic change in the UK, 1960–2012. **a**, Long-term changes in air temperature and precipitation are the differences between the 1960 and 2012 monthly means of these variables, derived from a regression fitted through each monthly time series. Error bars indicate the standard deviation of linearly detrended climatological data, as an indication of inter-annual variation around each trend. **b, c**, Examples of spatial variation in the extent of long-term climatic changes are shown for March air temperatures (**b**) and February precipitation (**c**).

delays with rising precipitation was greatest: the average upper limit of CSP_{precip} exceeded the lower limit (1.4 days per mm and -0.4 days per mm, respectively; Fig. 3b and Extended Data Table 2). The upper limit of CSP_{precip} was more variable among species than among sites (species and site level random-effects variances in the seasonal timing–precipitation slopes were 1.9 and 1.2, respectively). The fitted climate–phenology model was better supported by the data than a year-only model with the same random effects structure (ΔAIC (Akaike’s information criterion) 293,516). This indicates the presence of real associations between climate and seasonality, rather than purely spurious correlations resulting from shared temporal trends. Average sensitivity to temperature was very similar in the model that included marine plankton data, but excluded precipitation effects (see Supplementary Discussion and Extended Data Fig. 1).

Taxonomic and trophic group sensitivity

We tested the hypothesis that the limits of seasonal climate sensitivity differ coherently among taxonomic groups by including a fixed-effect interaction between taxonomic group and each climatic variable (Fig. 4 and Extended Data Table 2). The lower limit of CSP_{temp} was negative for all groups (‘earliness’ with warming), the strongest responses being found for plants, freshwater phytoplankton, insects and amphibians (4.3, 4.1, 3.7 and 3.4 days earlier per $^{\circ}C$, respectively). The upper limits of CSP_{temp} indicated that freshwater phytoplankton and mammals experienced the greatest phenological delays with seasonal warming (2.9 and 2.0 days later per $^{\circ}C$, respectively) but that plants showed little evidence of such delays. The strongest phenological delays with rising seasonal precipitation were found for freshwater phytoplankton and insects (2.5 and 2.2 days later per mm, respectively), while freshwater phytoplankton also exhibited the strongest phenological advances with rising precipitation during other seasonal windows (1.1 days earlier per mm). Average temperature and precipitation responses were consistent with CSP type II in most cases. There was considerable within-group variability in sensitivity.

We examined trophic-level differences in climate sensitivity by including trophic level in interaction with each climate variable in the global model. The lower limit of CSP_{temp} showed greater systematic variation than the upper limit among trophic levels (Fig. 3c, e). The tendency towards ‘earliness’ with seasonal warming was strongest at lower

trophic levels (-4.1 , -3.7 and -1.9 days per $^{\circ}C$ for primary producers, primary consumers and secondary consumers, respectively; Extended Data Table 2), consistent with observations of more rapid phenological changes at lower trophic levels, in the UK⁸. Conversely, the lower limit of CSP_{precip} varied less than the upper limit among trophic levels (Fig. 3d, f). The tendency for seasonal events to be later with higher seasonal precipitation was greater for primary producers and primary consumers (1.8 and 2.2 days per mm on average, respectively) than for secondary consumers (1.0 days per mm). Variations in climate

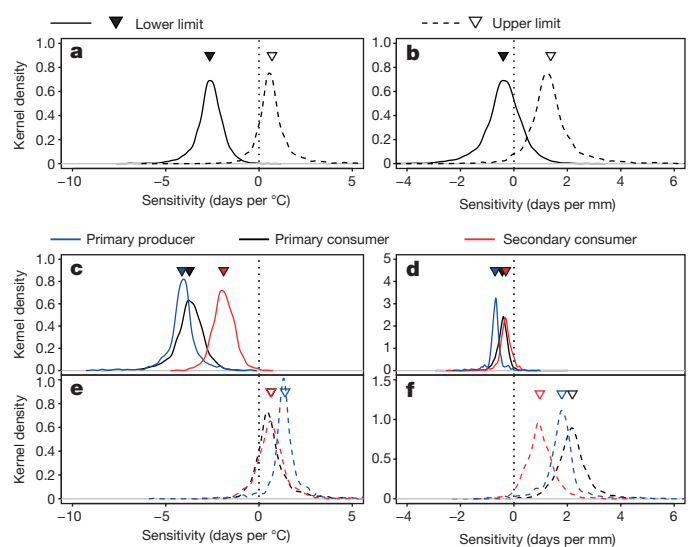


Figure 3 | Upper and lower limits of phenological climate sensitivity.

Sensitivity is the slope of the relationship between seasonal timing (day of year) and climatic variables. **a, b**, All-taxa upper and lower limits in sensitivity to temperature ($^{\circ}C$; **a**) and precipitation (mm per day; **b**) are summarized. **c–f**, Lower (c, d) and upper (e, f) limits of sensitivity to temperature (c, e) and precipitation (d, f) are shown by trophic level. Inverted triangles indicate average sensitivity. Curves are kernel density plots: estimates of the probability density distribution of species-level climate sensitivity (that is, the relative likelihood of different levels of climate sensitivity within each species group) ($n = 370,725$).

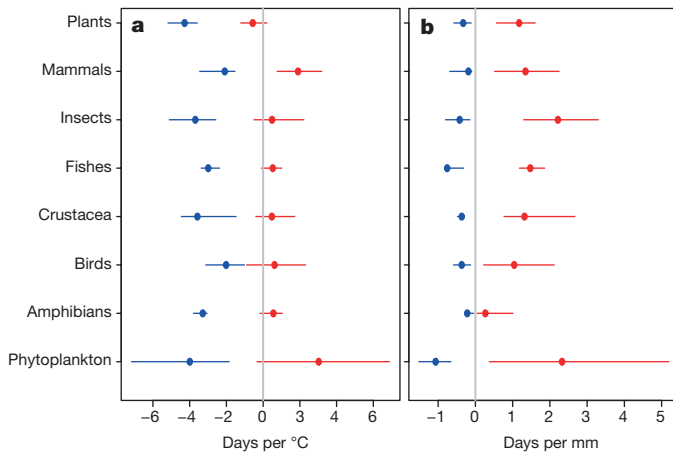


Figure 4 | Upper and lower limits of phenological climate sensitivity for broad taxonomic groups. **a**, **b**, Lower (blue) and upper (red) limits of the sensitivity of phenological events to seasonal temperature (**a**) and precipitation (**b**) change are shown. Coloured circles indicate the median response, and bars show the 5th–95th percentile responses for each group. Sensitivity is quantified by summarizing the species-level (random effects) responses from a mixed effects model including data for all taxa, and with taxonomic group as a fixed effect ($n = 370,725$).

sensitivity were described more parsimoniously by taxonomic groups than by trophic levels (AICs of taxonomic and trophic-level models 3,237,611 and 3,238,061, respectively).

The results were affected little when we analysed only pre- and post-1980 data, to minimize among-group variation in time series length, and after Monte Carlo re-sampling to assess the potential effects of taxonomic bias (see Supplementary Discussion and Extended Data Figs 2–4). The same qualitative trophic-level differences in climate sensitivity were apparent when we included marine plankton data in a temperature-only LME model (see Supplementary Discussion and Extended Data Fig. 1). In contrast to trophic-level differences in the magnitude of sensitivity, there was little evidence of similar variation in the seasonal timing of climate sensitivity (see Supplementary Discussion and Extended Data Figs 5–7).

Estimating future change

Overall, net phenological responses to climatic change combine potentially opposing responses to conditions in different seasonal periods. We estimated net responses by the 2050s by applying our fitted models to UKCP09 probabilistic projections (bias-corrected relative to a 1961–1990 baseline) of temperature and precipitation change under low-, medium- and high-emissions scenarios. Rather than predicting the absolute timing of future phenological events, we contrasted possible changes in seasonal timing among organism groups based upon established climate scenarios and contemporary patterns of climate sensitivity. Estimated average phenological changes were less for primary producers and secondary consumers than for primary consumers (Fig. 5a). This occurred because, averaged across species, the opposing climate responses of primary producers and secondary consumers are more similar in magnitude than are those of primary consumers (Fig. 3), effectively cancelling each other out. Our models suggest greater average advances for crustacea, fish and insects than for other groups, such as freshwater phytoplankton, birds and mammals (Fig. 5b). However, response variation is high for crustacea (Fig. 5b).

Discussion

In the UK, phenological climate sensitivity varies greatly, suggesting that it is influenced by locally varying, non-climatic drivers such as population structure¹⁸, resource availability¹⁹ and adaptation²⁰. This is relevant to the use of phenological change as a tangible climate change

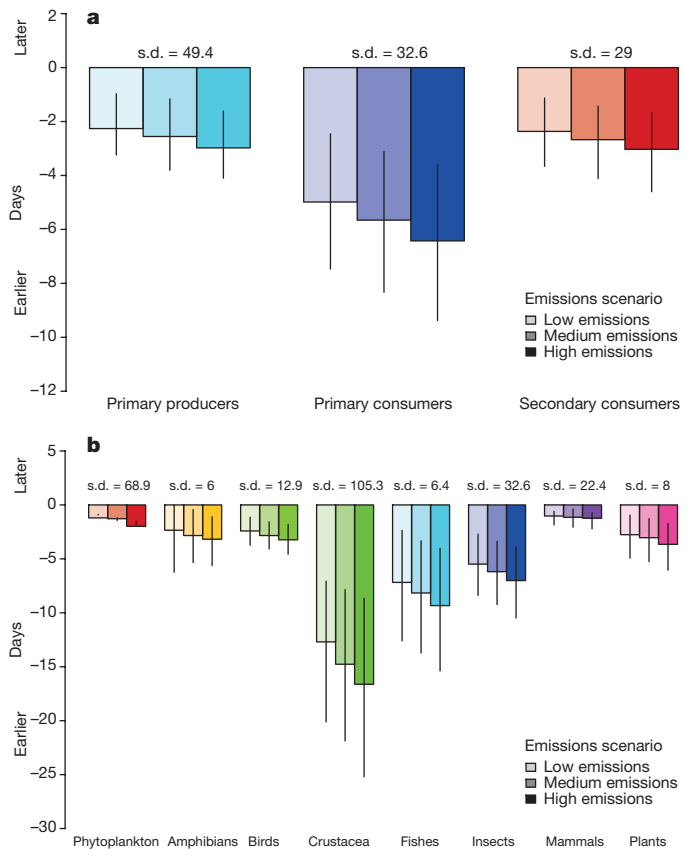


Figure 5 | Estimated phenological shifts by the 2050s. **a**, **b**, Modelled responses to projected temperature and precipitation change, assuming contemporary climate sensitivity, for trophic levels (**a**) and taxonomic groups (**b**). Projected median shifts in seasonal timing are shown. Change estimates are based on low-, medium- and high-emissions climate scenarios. Bars represent median responses to 50th percentile climate change projections under each scenario; extremes of whiskers represent median responses to the 10th and 90th percentile projected climatic changes under each scenario. Standard deviations indicate variation in projected responses for each group under the 50th percentile of the medium-emissions scenario.

indicator^{1,21}. Mediators of phenological climate sensitivity are known only locally for some of the groups in our data set (for example, nutrient availability for freshwater phytoplankton)²². However, for others, the climate sensitivity of different biological traits is known to be mediated by alternative drivers^{23,24}. High climate response variability necessitates wide site and species coverage in long-term monitoring schemes designed to develop robust aggregate indicators of change²¹. As climatic conditions are more spatially variable across broader geographic domains, site-level replication of phenological monitoring is particularly important when interpreting phenology at continental-to-global scales. In the UK, average responses for fish and insects appear to be sensitive indicators of climate effects. These groups show consistently strong phenological advances with seasonal warming, and only weak opposing responses, resulting in relatively large (net) changes in seasonal timing. Interpretation of phenological changes for other groups is more complex. For example, freshwater phytoplankton show strong evidence of opposing phenological responses to climatic variation at different times of year and these responses are near-equivalent in magnitude, such that estimated net changes are negligible. This finding emphasizes that long-term observations represent the net effect of potentially opposing biological responses²⁵. To fully capitalize on the indicator potential of phenological change, we must advance our mechanistic understanding of responses to potentially opposing climatic and non-climatic drivers.

Despite this variability, we identified coherent patterns in climate sensitivity among the idiosyncratic responses of many wild plant and animal populations. We have shown that, on average, species in different trophic levels differ in the magnitude of seasonal climate sensitivity, but not in the time of year within which climatic change has its most pronounced effects. This may be a key mechanism underpinning observations of trophic level differences in phenological change in the UK⁸. Lower trophic levels demonstrated more pronounced variation in their sensitivity to changing temperature and precipitation at different times of year, and stronger phenological responses to climatic change during defined (taxon- and population-specific) seasonal periods.

In response to climatic changes projected for the 2050s, relative changes in seasonal timing are likely to be greatest for primary consumers, particularly in the terrestrial environment. The difference in magnitude between opposing climate responses is greatest for primary consumers, resulting in greater net change. Our approach makes the simplifying assumption that climatic change has an overriding influence upon seasonality. Nevertheless, our results suggest that systematic differences in climate sensitivity could result in widespread phenological desynchronization. However, factors that shape phenological climate responses introduce uncertainty into projections of future phenological change. These results should catalyse research to improve predictive capacity in the face of multiple environmental and demographic drivers that not only mediate rates of change, but might also confer resilience to desynchronization (for example, population density dependence²⁶, compensatory range shifts²⁷, and the formation of novel inter-specific interactions^{28,29}). These findings also underscore the importance of developing our capacity to manage ecosystems within a 'safe operating space' with respect to the likely impacts of projected climate change³⁰.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 12 June 2015; accepted 26 May 2016.

Published online 29 June 2016.

- IPCC. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* 1132 (Cambridge Univ. Press, 2014).
- Parmesan, C. & Yohe, G. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**, 37–42 (2003).
- Root, T. L. *et al.* Fingerprints of global warming on wild animals and plants. *Nature* **421**, 57–60 (2003).
- Both, C., van Asch, M., Bijlsma, R. G., van den Burg, A. B. & Visser, M. E. Climate change and unequal phenological changes across four trophic levels: constraints or adaptations? *J. Anim. Ecol.* **78**, 73–83 (2009).
- Visser, M. E., Holleman, L. J. M. & Gienapp, P. Shifts in caterpillar biomass phenology due to climate change and its impact on the breeding biology of an insectivorous bird. *Oecologia* **147**, 164–172 (2006).
- Burthe, S. *et al.* Phenological trends and trophic mismatch across multiple levels of a North Sea pelagic food web. *Mar. Ecol. Prog. Ser.* **454**, 119–133 (2012).
- Jonsson, T. & Setzer, M. A freshwater predator hit twice by the effects of warming across trophic levels. *Nat. Commun.* **6**, 5992 (2015).
- Thackeray, S. J. *et al.* Trophic level asynchrony in rates of phenological change for marine, freshwater and terrestrial environments. *Glob. Change Biol.* **16**, 3304–3313 (2010).
- Visser, M. E. & Both, C. Shifts in phenology due to global climate change: the need for a yardstick. *Proc. R. Soc. Lond. B* **272**, 2561–2569 (2005).
- Walpole, M. *et al.* Ecology. Tracking progress toward the 2010 biodiversity target and beyond. *Science* **325**, 1503–1504 (2009).
- Butchart, S. H. M. *et al.* Global biodiversity: indicators of recent declines. *Science* **328**, 1164–1168 (2010).
- Williams, S. E., Shoo, L. P., Isaac, J. L., Hoffmann, A. A. & Langham, G. Towards an integrated framework for assessing the vulnerability of species to climate change. *PLoS Biol.* **6**, e325 (2008).
- Post, E. & Forchhammer, M. C. Climate change reduces reproductive success of an Arctic herbivore through trophic mismatch. *Philos. Trans. R. Soc. B Biol. Sci.* **363**, 2367–2373 (2008).

- Thackeray, S. J., Jones, I. D. & Maberly, S. C. Long-term change in the phenology of spring phytoplankton: species-specific responses to nutrient enrichment and climatic change. *J. Ecol.* **96**, 523–535 (2008).
- Doi, H., Gordo, O. & Katano, I. Heterogeneous intra-annual climatic changes drive different phenological responses at two trophic levels. *Clim. Res.* **36**, 181–190 (2008).
- Visser, M. E., van Noordwijk, A. J., Tinbergen, J. M. & Lessells, C. M. Warmer springs lead to mistimed reproduction in great tits (*Parus major*). *Proc. R. Soc. Lond. B* **265**, 1867–1870 (1998).
- van de Pol, M. & Cockburn, A. Identifying the critical climatic time window that affects trait expression. *Am. Nat.* **177**, 698–707 (2011).
- Ohlberger, J., Thackeray, S., Winfield, I., Maberly, S. & Vøllestad, L. When phenology matters: age-size truncation alters population response to trophic mismatch. *Proc. R. Soc. Lond. B* **281**, 20140938 (2014).
- Thackeray, S. J., Henrys, P. A., Jones, I. D. & Feuchtmayr, H. Eight decades of phenological change for a freshwater cladoceran: what are the consequences of our definition of seasonal timing? *Freshw. Biol.* **57**, 345–359 (2012).
- Phillimore, A. B., Hadfield, J. D., Jones, O. R. & Smithers, R. J. Differences in spawning date between populations of common frog reveal local adaptation. *Proc. Natl Acad. Sci. USA* **107**, 8292–8297 (2010); correction **109**, 5134 (2012).
- Amano, T., Smithers, R. J., Sparks, T. H. & Sutherland, W. J. A 250-year index of first flowering dates and its response to temperature changes. *Proc. R. Soc. Lond. B* **277**, 2451–2457 (2010).
- Feuchtmayr, H. *et al.* Spring phytoplankton phenology — are patterns and drivers of change consistent among lakes in the same climatological region? *Freshw. Biol.* **57**, 331–344 (2012).
- Nussey, D. H., Clutton-Brock, T. H., Albon, S. D., Pemberton, J. & Kruuk, L. E. B. Constraints on plastic responses to climate variation in red deer. *Biol. Lett.* **1**, 457–460 (2005).
- Van Emden, H. F. & Harrington, R. *Aphids as Crop Pests*. 717 (CABI, 2007).
- Cook, B. I., Wolkovich, E. M. & Parmesan, C. Divergent responses to spring and winter warming drive community level flowering trends. *Proc. Natl Acad. Sci. USA* **109**, 9000–9005 (2012).
- Reed, T. E., Grøtan, V., Jenouvrier, S., Sæther, B.-E. & Visser, M. E. Population growth in a wild bird is buffered against phenological mismatch. *Science* **340**, 488–491 (2013).
- Amano, T. *et al.* Links between plant species' spatial and temporal responses to a warming climate. *Proc. R. Soc. Lond. B* **281**, 20133017 (2014).
- Miller-Rushing, A. J., Hoyer, T. T., Inouye, D. W. & Post, E. The effects of phenological mismatches on demography. *Philos. Trans. R. Soc. B* **365**, 3177–3186 (2010).
- Nakazawa, T. & Doi, H. A perspective on match/mismatch of phenology in community contexts. *Oikos* **121**, 489–495 (2012).
- Scheffer, M. *et al.* Creating a safe operating space for iconic ecosystems. *Science* **347**, 1317–1319 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was funded by Natural Environment Research Council (NERC) grant NE/J02080X/1. We thank O. Mountford for assigning species traits for plants, H. Feuchtmayr for extracting plankton data for analysis and N. Dodd for air and water temperature data from the Tarland Burn. We also thank P. Verrier, the staff and many volunteers and contributors, including Science and Advice for Scottish Agriculture, to the Rothamsted Insect Survey (RIS) over the last half century. The RIS is a National Capability strategically funded by BBSRC. The consortium represented by the authorship list hold long-term data that represent a considerable investment in scientific endeavour. Whilst we are committed to sharing these data for scientific research, users are requested to collaborate before publication of these data to ensure accurate biological interpretation.

Author Contributions S.J.T. and S.W. conceived and coordinated the study and led writing of the manuscript. P.A.H. developed the analysis routine and wrote statistical code to be applied to all data sets. D.H. extracted all climatic and sea surface temperature data. I.D.J. and E.B.M. calculated water temperatures for lakes and streams, respectively. S.J.T., J.R.B., M.S.B., S.B., P.H., T.T.H., D.G.J., D.I.L., E.B.M. and D.M. led analysis of specific data sets using code from P.A.H. S.A., P.J.B., T.M.B., L.C., T.H.C.-B., C.D., M.E., J.M.E., S.J.G.H., R.H., J.W.P.-H., L.E.B.K., J.M.P., T.H.S., P.M.T., I.W. and I.J.W. derived phenological data for analysis, advised on interpretation, and assisted in assigning species traits. All co-authors commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.J.T. (sjtr@ceh.ac.uk).

Reviewer Information Nature thanks D. Inouye, M. Visser and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Data sets. We integrated data from many major UK biological monitoring schemes (Supplementary Table 1), resulting in 10,003 long-term (at least 20 years between 1960 and 2012) phenological series for 812 marine, freshwater and terrestrial taxa. The amassed data sets included records for plants, phytoplankton, zooplankton, insects, amphibians, fish, mammals and birds (379,081 individual phenological observations). For each study we used a single population-level phenological measure per year (Supplementary Table 1). Because the sampling resolution for the marine plankton data was monthly, before analysis we re-scaled these data into units of days. Trophic level, taxonomic class and environmental affinity were assigned to each taxon, to permit analyses of correlations between these attributes and climate sensitivity.

Daily air temperature and precipitation data were extracted from the Met Office National Climate Information Centre (NCIC) 5-km-resolution gridded data set³¹ for the spatial locations of all biological monitoring sites across the UK land surface. If available, recorded water temperatures from the same sites were used in place of air temperatures for phenological time series representing obligate aquatic taxa (freshwater plankton and fish). Water temperatures were interpolated onto a daily time-step before analysis³². If these data were not available, daily water temperature data were estimated from air temperatures using a fitted empirical site-specific relationship between air and water temperature. For the sea trout (*Salmo trutta*) data, an existing linear relationship³³ was used, while for the Atlantic salmon (*Salmo salar*) data, a nonlinear relationship³⁴ was calculated for a nearby river, the Tarland Burn, and applied to air temperatures from the sampling site. For the marine plankton, mean monthly sea surface temperatures were extracted from the Met Office Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) data set³⁵ for each of the Standard Areas³⁶ in which phenological data were available. Precipitation data were not available for marine Standard Areas.

Statistics. Our analysis was conducted in two distinct phases (see Supplementary Notes). First, the CSP for each phenological series was estimated using generalized linear models to quantify associations between the timing of seasonal events and mean temperature and precipitation (within defined seasonal time windows) at the same location. Second, the phenological time series were aggregated and a single LME model was run, capturing upper and lower limits of climate sensitivity across many species. CSPs for precipitation were not estimated for marine plankton data (see above), so the second-phase LME models were run twice: once to examine correlations with temperature and precipitation for all but the marine plankton phenological series (9,800 series), and once to examine only correlations with temperature for the whole data set (10,003 series).

Phase 1: Estimating CSPs for each time series. We used consistent methods to 'screen' all phenological events with respect to their climate sensitivity, finding periods of the year in which temperature and precipitation had the most positive and negative correlations with seasonal timing (the upper and lower limits of climate sensitivity). This approach was flexible enough to detect when these limits represented opposing correlations between temperature or precipitation and seasonality, depending upon the seasonal timing of climatic change; for example, spring warming may advance budburst, but winter warming may delay it³⁷ (Fig. 1; CSP type II). This approach could also detect when the direction of the correlation between climatic variables and seasonal timing was consistent irrespective of the seasonal timing of climatic change, with only the magnitude of the correlation varying between the limits of the CSP (Fig. 1; CSP types I and III).

For each phenological time series, we calculated the day of the year by which 95% of the recorded seasonal events had occurred (doy₉₅). Inter-annual variations in seasonal timing were statistically modelled as a function of daily mean temperatures on doy₉₅ each year. Then, a series of 365 statistical models was run that instead used daily mean temperatures on doy₉₅-1 to doy₉₅-365 as predictors. Slope coefficients and R^2 values for the temperature terms in these models were collated, capturing seasonal variations in the sign and magnitude of the phenology-temperature relationship (that is, the CSP; Fig. 1). Generalized linear models (GLMs) were used.

For two data sets (BTO Nest Record Scheme and PTES National Dormouse Monitoring Scheme; Supplementary Table 1) we modified the above analytical framework. In both of these schemes, the precise location of the biological observations changed among years (compared with other schemes in which monitoring sites were static over time). We extracted matching climatic data for each specific location in each year, as for all other schemes, but then grouped the phenological and climatic data at county level (mean area 3,440 km²). Then, for each taxon in each county we used the fixed-effect slope parameters and R^2 values from a series of LME models, instead of GLMs, as a basis for estimating CSPs. In these models, we included fixed effects of temperature on doy₉₅ to doy₉₅-365 as before, and included a year random effect to account for replicate phenological records for each taxon in each county in each year. For the SAHFOS marine plankton data

set, we modified our iterative approach to analyse seasonal timing-temperature relationships at monthly, instead of daily, time steps (the temporal resolution of the sea surface temperature data).

As a final step in estimating the CSP for each series, temporal variation in the sign and magnitude of the seasonal timing-temperature correlation was itself modelled (Extended Data Fig. 8). This was done by fitting generalized additive models (GAMs, gamma error distribution) to the time series of slope coefficients and R^2 values from the models described above. By smoothing these time series, the GAMs identified periods of the year in which slope coefficients were consistently negative (that is, warming advanced seasonal timing), or consistently positive (that is, warming delayed seasonal timing), and during which the climate-phenology models generating the slope estimates had a their highest goodness-of-fit.

Seasonal windows in which the upper and lower limits of temperature sensitivity occurred were identified as periods during which: 1) the 95% confidence interval for the GAM fitted to the slope coefficients surpassed the limits of the 2.5th and 97.5th percentiles of the original slope coefficients; and 2) the 95% confidence interval for the GAM fitted to the R^2 values surpassed the 97.5th percentile of the original R^2 values. This ensured that seasonal windows were defined by periods combining the greatest climate effect size and relatively strong predictive power (determined by R^2). Using this framework, we identified the lower limit of CSP_{temp}: the period of the year in which an advancing effect of increasing temperature upon seasonal timing was most likely. This was estimated by determining when the 95% confidence interval of the GAM intersected the lower percentile of the seasonal timing-temperature slope coefficients, by tracking the most negative coefficients (Extended Data Fig. 8). In addition, we identified the upper limit of CSP_{temp} by determining when the 95% confidence interval of the GAM intersected the upper percentile of the seasonal timing-temperature slope coefficients, by tracking the most positive (or least negative) coefficients. Excluding the marine plankton data, the whole modelling process was repeated with precipitation as a predictor instead of air temperature, culminating in the estimation of seasonal periods capturing the limits of phenological responses to changing precipitation.

After this process, temperature and precipitation were each averaged within the two seasonal windows in which the limits of phenological sensitivity occurred. With the exception of the marine plankton data, the final seasonal timing-climate model for each series was then fitted using a GLM with gamma error distribution including four predictors: inter-annual variations in 1) mean temperature during the period at the lower limit of CSP_{temp}, 2) mean temperature during the period at the upper limit of CSP_{temp}, 3) mean precipitation during the period at the lower limit of CSP_{precip}, 4) mean precipitation during the period at the upper limit of CSP_{precip}. For the marine plankton data, only the first two terms were fitted. For the BTO Nest Record and PTES National Dormouse Monitoring Scheme data sets we implemented these final models in a mixed effects framework with a random effect of year, as before. Therefore, although we modelled changes in statistical parameters (which are not estimated without error) to identify seasonal periods, this step was used only to find the original climatic data to be used in subsequent modelling. Inferences were not, therefore, directly based upon statistical modelling of uncertain parameter estimates. We categorized the results of all 10,003 CSPs according to three broad response types (CSP types I-III; Fig. 1), and retained P values for each fitted model term to infer which of the modelled climatic effects were statistically significant. We examined the evidence for trophic-level differences in the mean seasonal timing of climate sensitivity by modelling the relationship between the start date, end date and duration of the seasonal windows capturing the upper and lower limits of phenological sensitivity to temperature and rainfall as a function of trophic level (fixed effect), with random effects of phenological metric, within species, within site. Analyses were conducted using the *base*, *mgcv* and *lme4* packages in R (refs 38-40).

Phase 2: Global models of phenological climate sensitivity. We estimated the upper and lower limits of phenological climate sensitivity at a multi-species scale by matching each phenological series with data on mean temperature and precipitation, during the seasonal windows characterizing the CSP for that series (Phase 1, above). We aggregated all 10,003 of these matched phenology-climate data sets. To quantify the average, multi-species, upper and lower limits of climate sensitivity we constructed an LME model in which phenology (day of year) was modelled as a function of mean temperature and precipitation within the seasonal windows of the amassed CSPs (fixed effects) with random effects of phenological metric, within species, within site. These random effects were necessary because our data could not be considered independent. The timings of events are more likely to be similar for the same species than for different species; the same is true for different sites and the phenological metric-types used to describe the events (for example, first flight time or seasonal peak abundance). Random slopes and intercepts were allowed to ensure that each phenological event, for a species at a site, was allowed a different rate of climate response.

For some species, more than one phenological event was recorded in the same year, at the same site. For example, butterflies may have more than one flight period in the same year, and plankton populations may be characterized by more than one seasonal abundance peak. As climate responses are unlikely to be the same for the first event of the year and subsequent events, we introduced a voltinism factor in the analysis. This allowed us to distinguish between data representing the first or only events of each year (for example, a spring plankton bloom or butterfly generation) and second events in each year (for example, the subsequent summer plankton bloom or butterfly generation). This distinction captured all possibilities within our data set.

For site i , species j , voltmetric k (where voltmetric is a unique combination of voltinism class and the metric-type used to identify the event), the corresponding day of year (DOY) of a particular seasonal event is modelled as:

$$\text{DOY}_{ijk} = \alpha_0 + \beta_1 \text{TL}_{ijk} + \beta_2 \text{TU}_{ijk} + \beta_3 \text{PL}_{ijk} + \beta_4 \text{PU}_{ijk} + \varepsilon_{ijk}$$

where $\varepsilon_{ijk} \sim N(0, \sigma^2)$ and the model includes temperature at the upper limit of each CSP (TU), temperature at the lower limit of each CSP (TL), precipitation at the upper limit of each CSP (PU) and precipitation at the lower limit of each CSP (PL). Owing to the non-independence within the data, we allow the intercepts and coefficients corresponding to all four covariates to vary by site, species and voltmetric. Preserving the natural nesting of a metric for a species at a particular site, this gives:

$$\alpha_0 = \gamma_0 + \mu_{0,ijk} + \mu_{0,ij,k} + \mu_{0,i,jk}$$

$$\beta_1 = \gamma_1 + \mu_{1,ijk} + \mu_{1,ij,k} + \mu_{1,i,jk}$$

$$\beta_2 = \gamma_2 + \mu_{2,ijk} + \mu_{2,ij,k} + \mu_{2,i,jk}$$

$$\beta_3 = \gamma_3 + \mu_{3,ijk} + \mu_{3,ij,k} + \mu_{3,i,jk}$$

$$\beta_4 = \gamma_4 + \mu_{4,ijk} + \mu_{4,ij,k} + \mu_{4,i,jk}$$

where each of the μ terms is a normally distributed random effect.

This nesting of random effects is most conservative in terms of inference at the global level and is as flexible as possible, allowing each time series to have its own set of model parameters. This permits a high degree of biological realism because each distinct phenological event, for a given species, at a given site, is permitted to have a different slope for the effects of temperature and precipitation (that is, a different climate sensitivity).

In this model framework we are specifically testing the null hypotheses that each of the climate variables shows no relation to the seasonal timing of biological events. Because of this, and the fact that each parameter is estimated directly, without distributional form assumed *a priori* or as the target distribution, we follow a frequentist approach to analysis. However, because the exact degrees of freedom cannot be evaluated when using restricted maximum likelihood, and hence no exact P values can be obtained, we present full summaries of all the parameters estimated at species level (as given by $\gamma + \mu_{ijk} + \mu_{ij,k}$, above). Approximate P values could be presented by taking conservative estimates of the degrees of freedom although, given the volume of data available, this will typically lead to the detection of many statistically significant results that may not be biologically significant. Examining the full range of estimated coefficients across the random effects levels ensures that we present the full range of variation around global parameters and can make more informed inferences. In this way we encourage the reader to interpret our results by using biological insight, not by depending upon P values alone.

To examine high-level differences in climate sensitivity among trophic levels and taxonomic groups, we re-fitted the LME model with these attributes as fixed-effect factors, interacting with the fixed-effect climate variables. The fixed-effect slopes from the resulting models allowed us to compare differences in phenological climate sensitivity among these broad organism groups, averaged across all taxa within each group. Supplementary Table 2 shows the number of

phenological series, sites and distinct taxa that contributed data to each of these groups. All models were run twice: once to examine correlations with both temperature and precipitation excluding marine plankton data (9,800 time series), and once to examine only temperature–phenology correlations for the whole data set (10,003 time series).

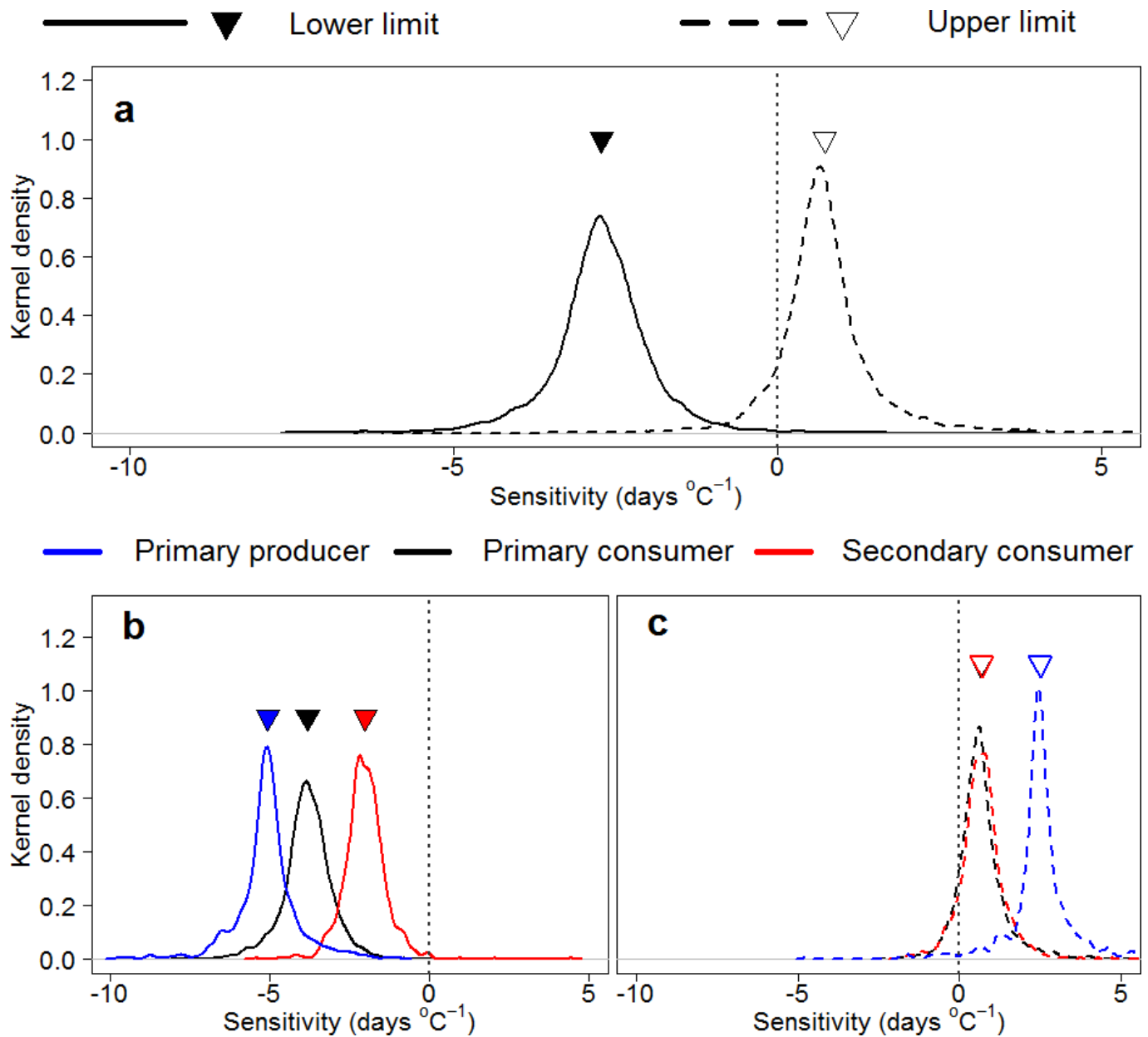
Potential biases. Data availability differed among taxonomic groups. To assess the extent to which mean responses were biased by data inequality we conducted Monte Carlo re-sampling, iteratively selecting 5, 20, 50 and 100 phenological series from each taxonomic group and re-fitting climate–phenology models with these sampled data sets. For taxonomic groups with fewer data than the larger sample sizes, we retained all available data (see Supplementary Discussion). This allowed us to compare taxonomic group and trophic level responses based on sampled and all data, to fully investigate potential bias.

Another potential bias in our analysis is that phenological time series length is variable, affecting the length of time over which climate–phenology correlations are assessed. In order to assess the extent to which differences in mean trophic level and taxonomic group responses are biased by variable time series length, we also re-fitted our models but based only on pre- and post-1980 data. All models were run in the *lme4* package in R (refs 38, 40).

Estimating future change. To estimate the potential future net effects of temperature and precipitation change, we compared predictions of seasonal timing under baseline conditions, and under established climate change scenarios. First, estimates of seasonal timing (day of year) were obtained for the same baseline period used in the UKCP09 projections (long term average 1961–1990), using modelled correlations between phenology, temperature and precipitation (from Phase 1). Having obtained these baseline estimates, we applied our models to projected changes in monthly temperature and precipitation for the 2050s (UK Climate Projections, UKCP09; <http://ukclimateprojections.metoffice.gov.uk/>). We used 10th, 50th and 90th percentile changes under low-, medium- and high-emissions scenarios (relative to the 1961–1990 baseline). The spatial location of each phenological series was matched to climate projection data for the 25×25 -km grid square in which it occurred, and temporally matched to climatic data from the months of year in which its respective climate sensitivity windows occurred. Relative changes in timing, in response to climatic change of the magnitude projected to occur by the 2050s, were summarized by trophic levels and taxonomic groups.

Code availability. R code to run the described analyses can be found on GitHub (https://github.com/NERC-CEH/Phenology_Climate).

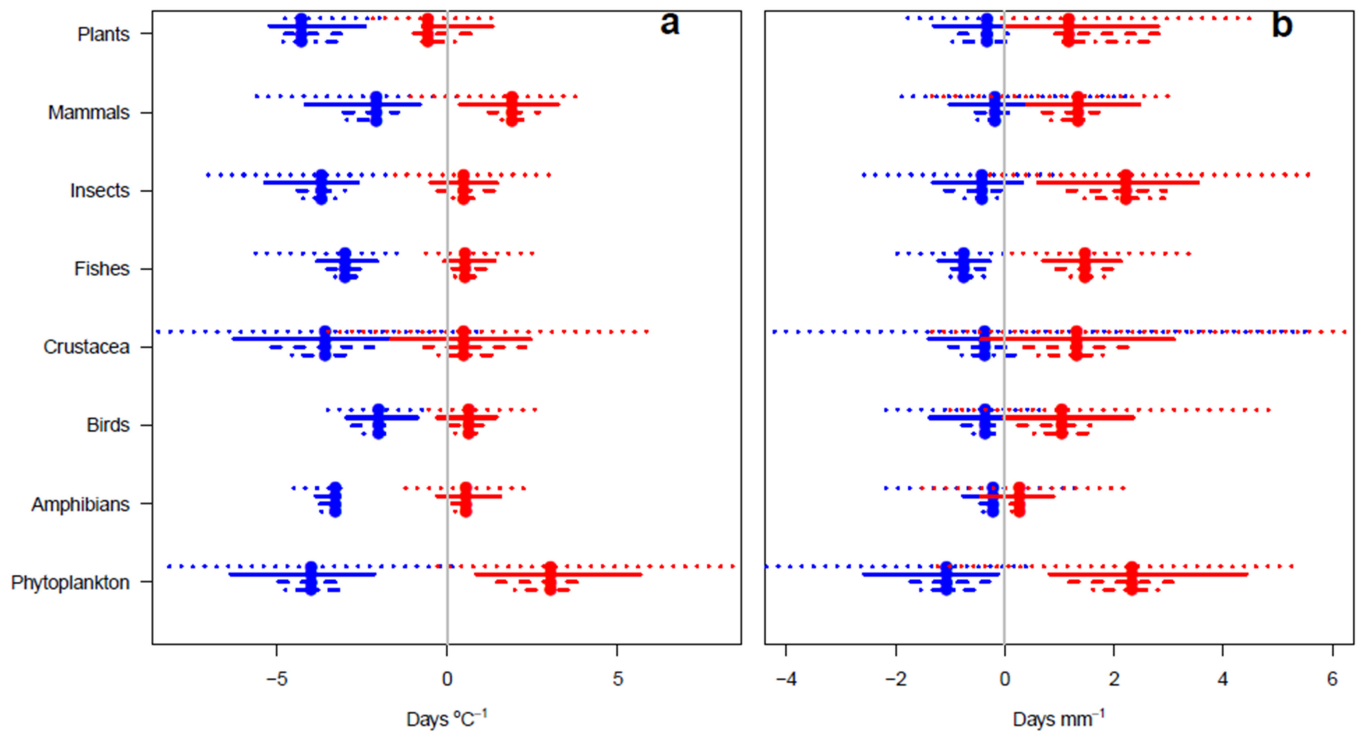
- Perry, M. & Hollis, D. The generation of monthly gridded data sets for a range of climatic variables over the UK. *Int. J. Climatol.* **25**, 1041–1054 (2005).
- Jones, I. D., Winfield, I. J. & Carse, F. Assessment of long-term changes in habitat availability for Arctic charr (*Salvelinus alpinus*) in a temperate lake using oxygen profiles and hydroacoustic surveys. *Freshw. Biol.* **53**, 393–402 (2008).
- Elliott, J. M. Numerical changes and population regulation in young migratory trout *Salmo trutta* in a Lake District stream, 1966–83. *J. Anim. Ecol.* **53**, 327–350 (1984).
- Mohseni, O., Stefan, H. G. & Erickson, T. R. A nonlinear regression model for weekly stream temperatures. *Wat. Resour. Res.* **34**, 2685–2692 (1998).
- Rayner, N. A. *et al.* Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.* **108**, 4407 (2003).
- Reid, P. C., Colebrook, J. M., Matthews, J. B. L. & Aiken, J. The Continuous Plankton Recorder: concepts and history, from Plankton Indicator to undulating recorders. *Prog. Oceanogr.* **58**, 117–173 (2003).
- Pope, K. S. *et al.* Detecting nonlinear response of spring phenology to climate change by Bayesian analysis. *Glob. Change Biol.* **19**, 1518–1525 (2013).
- R Development Core Team. *R: A Language and Environment for Statistical Computing* (2011).
- Wood, S. N. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *J. Am. Stat. Assoc.* **99**, 673–686 (2004).
- Bates, D., Maechler, M. & Bolker, B. *lme4: Linear Mixed-Effects Models using S4 Classes* (2011).



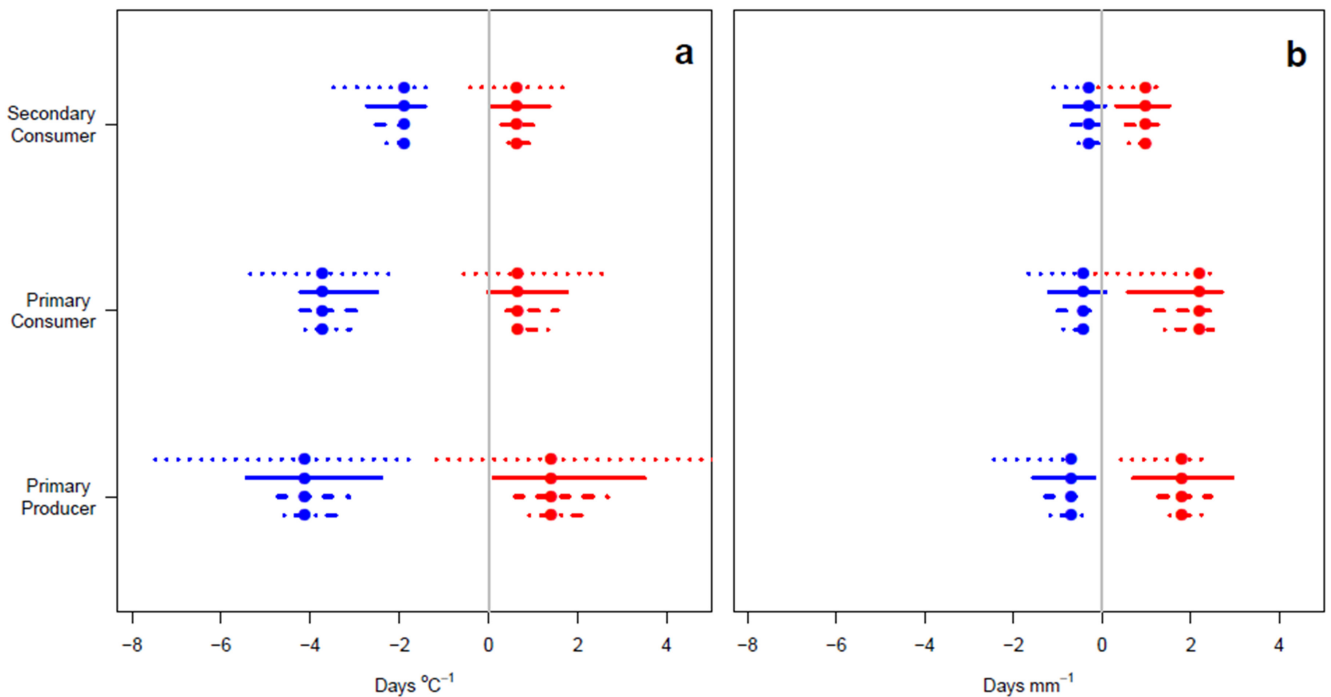
Extended Data Figure 1 | Limits of phenological temperature sensitivity inclusive of marine plankton data. a–c, Upper and lower limits of phenological temperature sensitivity are quantified as the slope of the relationship between seasonal timing (day of year) and temperature ($^\circ\text{C}$) variation within specific seasonal periods. Limits in temperature

sensitivity are shown for all taxa (a) and by trophic level (lower limit, b; upper limit, c). Inverted triangles indicate average sensitivity for all species in each group and curves are probability density plots of species-level variation in sensitivity ($n = 379,081$).

Taxonomic groups



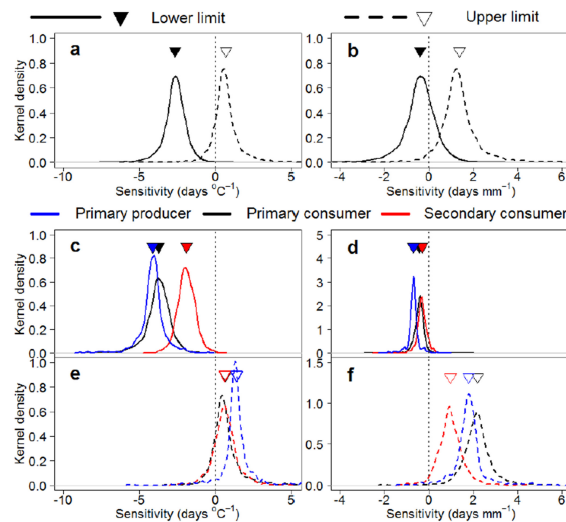
Trophic levels



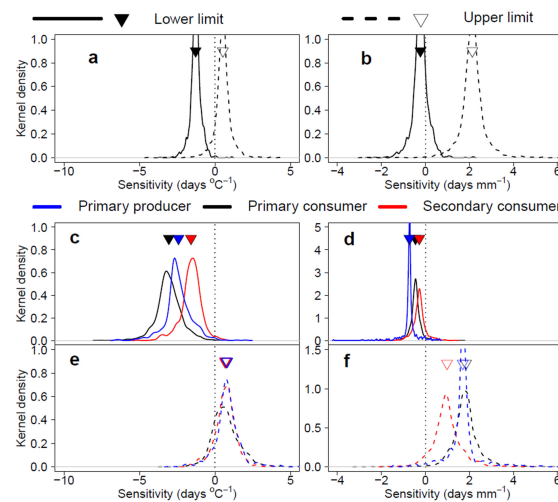
Extended Data Figure 2 | Limits of phenological climate sensitivity for taxonomic groups (top) and trophic levels (bottom), after Monte-Carlo resampling. a, b, Lower (blue) and upper (red) limits of the sensitivity of phenological events to changes in seasonal temperature (a) and precipitation (b). Coloured circles: responses based upon the full data set.

Bars: 2.5th–97.5th percentile responses for each group, based upon 100 draws from the full data set. Data were sampled so that 5, (dotted bar), 20 (solid bar), 50 (dashed bar) and 100 (dot-dashed bar) phenological time series were drawn from each taxonomic group ($n = 370,725$).

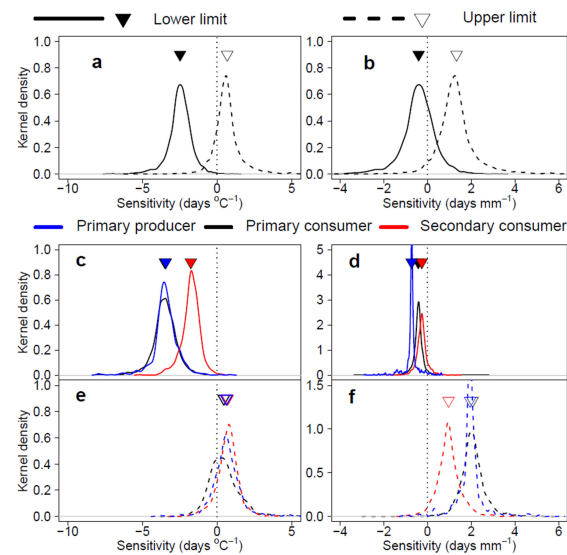
All data



Pre-1980 data



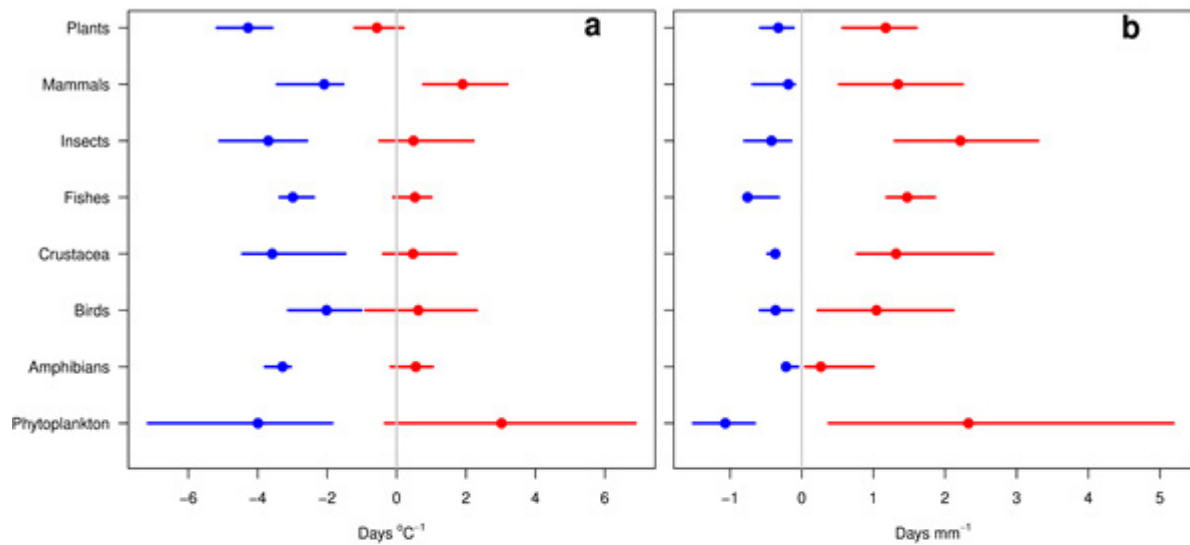
Post-1980 data



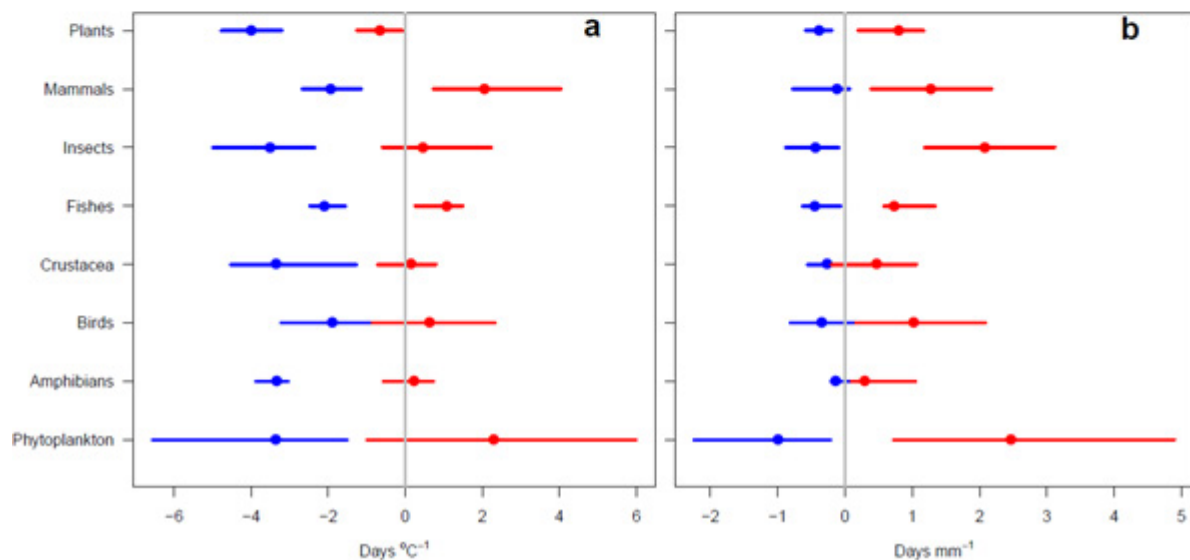
Extended Data Figure 3 | Climate sensitivities, based on different time periods. Top: all data; middle: pre-1980 data; bottom: post-1980 data. Sensitivity is the slope of the relationship between seasonal timing (day of year) and temperature (°C) or precipitation (mm per day). **a, b,** Limits of temperature (a) and precipitation (b) sensitivity are summarized for all taxa. **c–f,** Lower (c, d) and upper (e, f) limits of temperature (c, e)

and precipitation (d, f) sensitivity are shown by trophic level. Inverted triangles: average sensitivity for all species (a, b) or trophic levels (c–f). Curves, kernel density plots: probability density distributions of species-level climate sensitivity (that is, the relative likelihood of different climate sensitivities within each species group) ($n = 370,725$).

All data

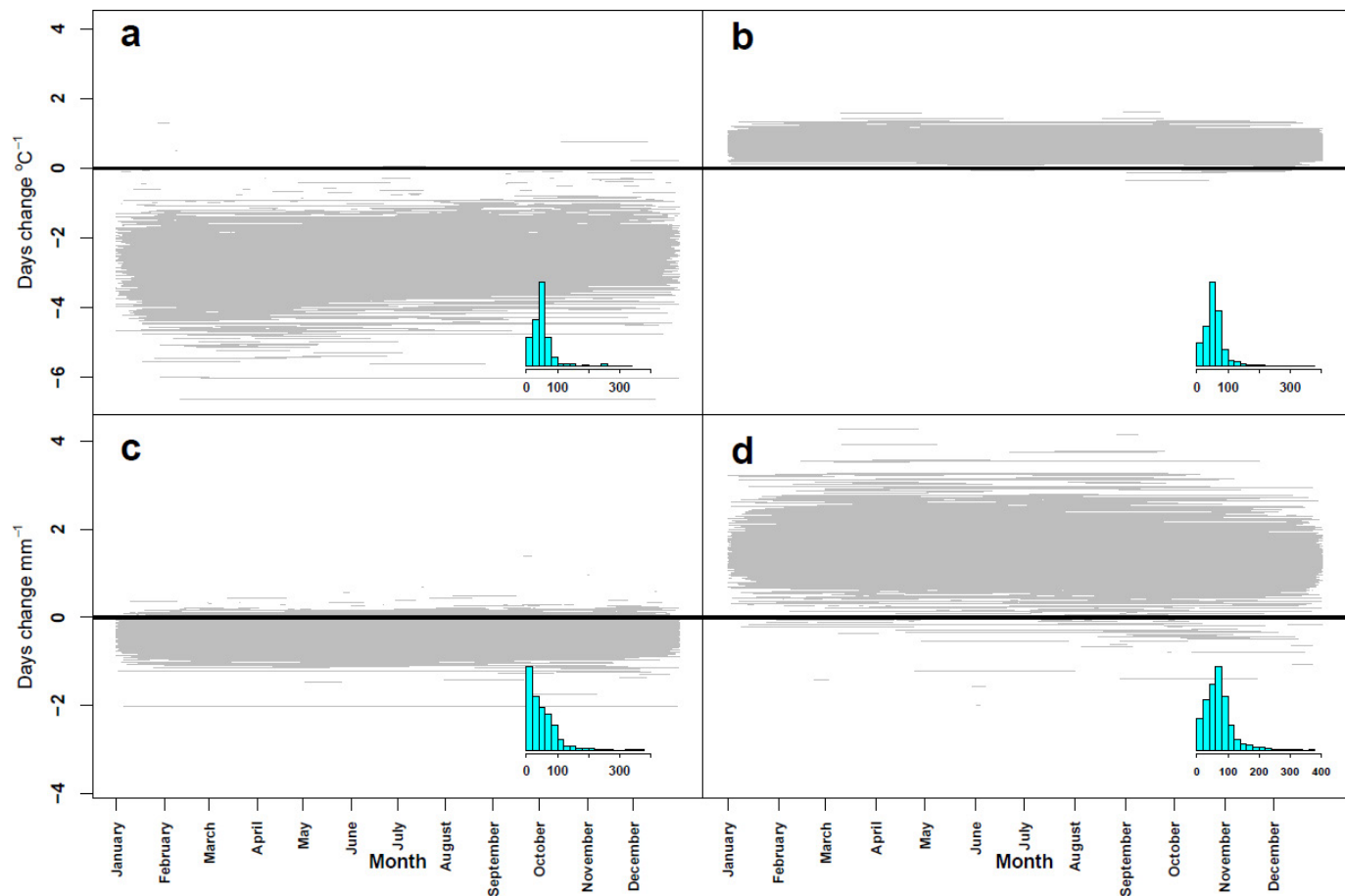


Post-1980 data



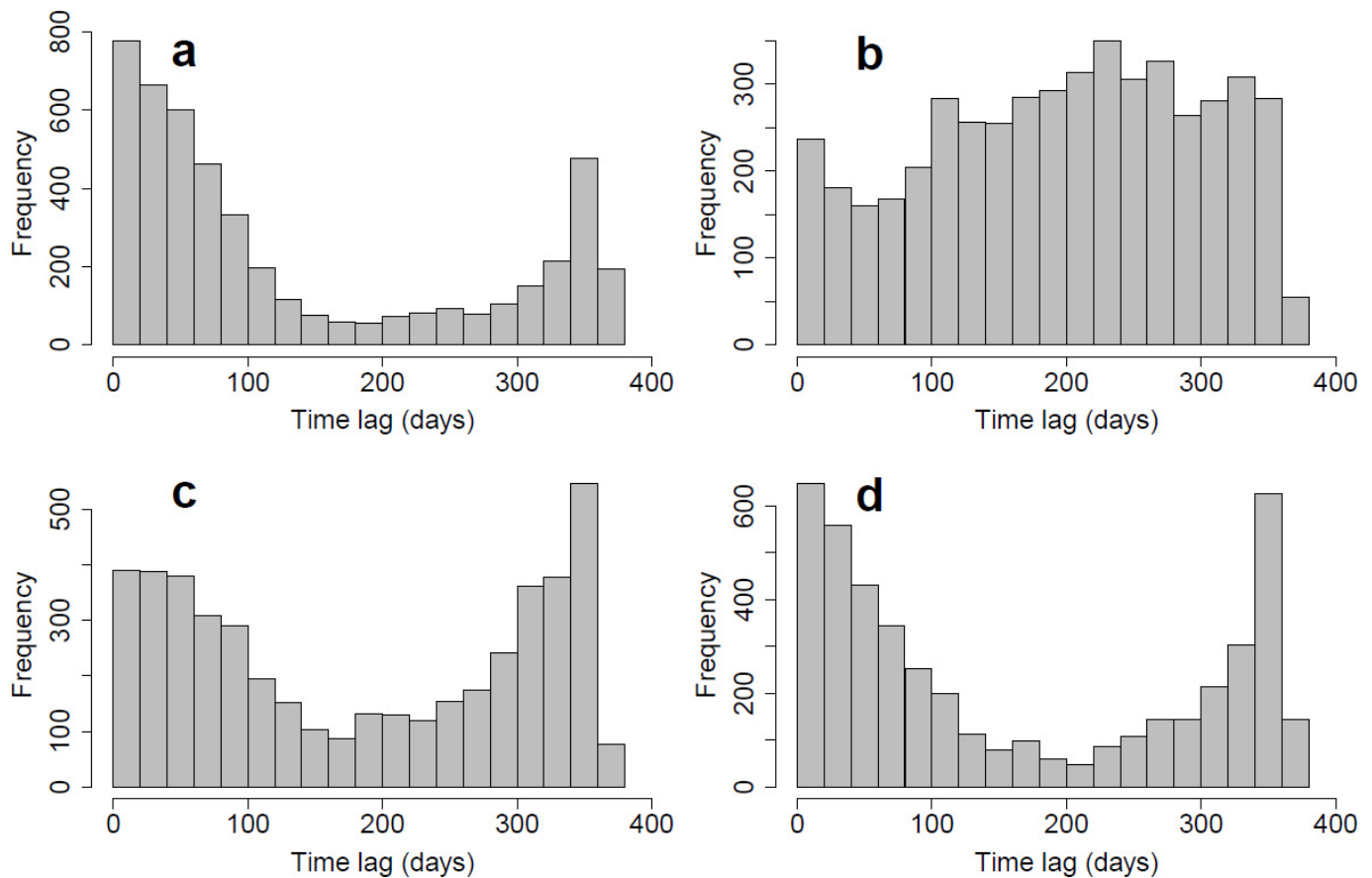
Extended Data Figure 4 | Limits of phenological climate sensitivity for broad taxonomic groups. Top, all data; bottom, post-1980 data only. **a, b**, Lower (blue) and upper (red) limits of the sensitivity of phenological events to seasonal temperature (**a**) and precipitation (**b**) change are shown. Coloured circles indicate the median response, and bars show the

5th–95th percentile responses for each group. Sensitivity is quantified by summarizing the species-level (random effects) responses from a mixed effects model including data for all taxa, and with taxonomic group as a fixed effect ($n = 370,725$).



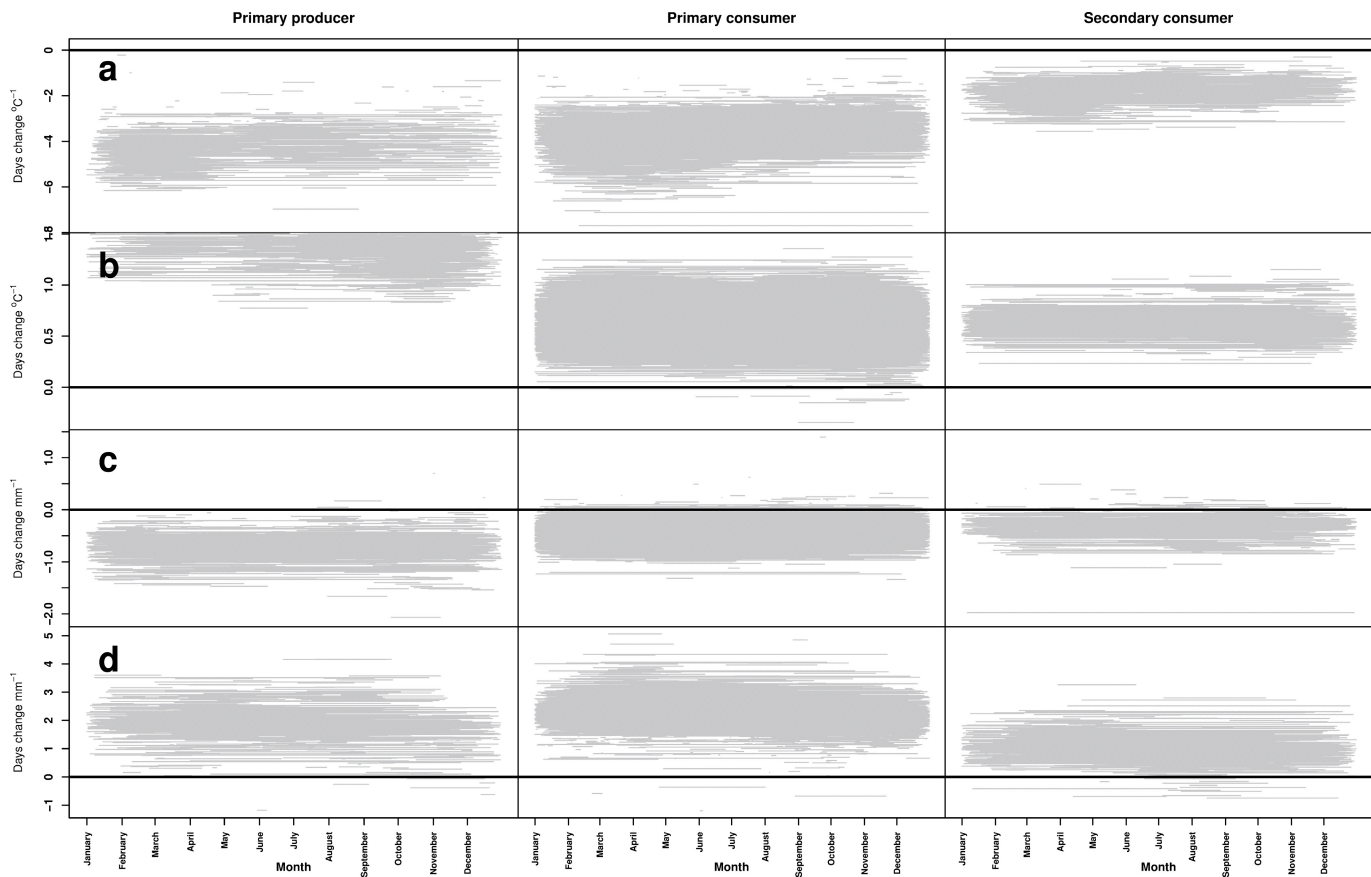
Extended Data Figure 5 | Seasonal windows for CSPs. a–d, Estimated climatic sensitivity at the lower (a, c) and upper (b, d) limits of CSPs for 10,003 phenological series. Grey lines are seasonal time periods (x -axis) within which climatic variables have their most positive or negative correlations with the seasonal timing of each phenological event. The y -axis

indicates the slope coefficient for each of these correlations; a measure of climate sensitivity (days change per $^{\circ}\text{C}$ or per mm). Shown are the lower and upper limits of CSP_{temp} (a and b, respectively) and the lower and upper limits of $\text{CSP}_{\text{precip}}$ (c and d, respectively). Inset histograms show seasonal time window length (days) ($n = 370,725$).

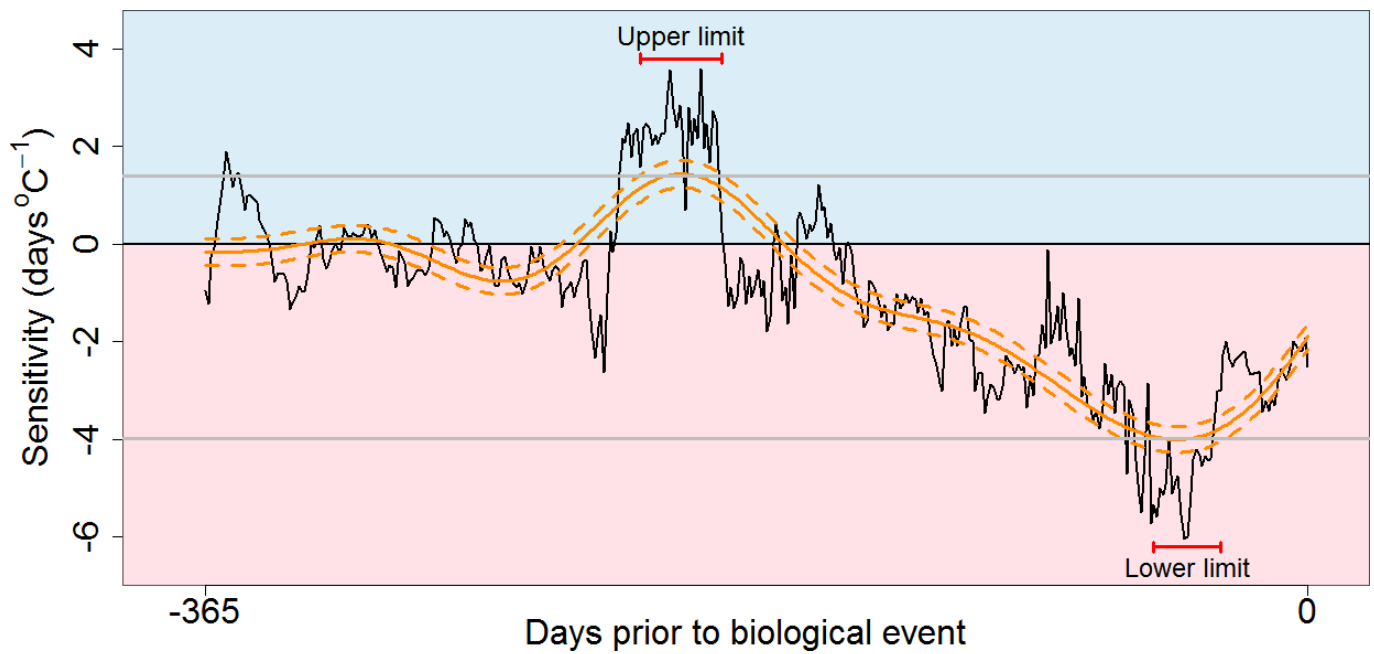


Extended Data Figure 6 | Time lags between phenological events and seasonal windows of climate sensitivity. a–d, Frequency histograms showing the time lag (in days) between the mean timing of each phenological event and the end of seasonal windows corresponding to the lower and upper limits of CSP_{temp} (a and b, respectively) and the

lower and upper limits of CSP_{precip} (c and d, respectively). Peaks at lags of around 1 year are where windows were identified that ended at the mean seasonal timing of an event, but in the previous year, owing to temporal autocorrelation in climate data ($n = 370,725$).



Extended Data Figure 7 | Seasonal windows for CSPs by trophic level. Estimated climatic sensitivity at the lower and upper limits of CSPs for taxa at each of three trophic levels. Formatting is as in Extended Data Fig. 5. **a–d**, Lower and upper limits of CSP_{temp} (**a** and **b**, respectively) and the lower and upper limits of $\text{CSP}_{\text{precip}}$ (**c** and **d**, respectively) ($n = 370,725$).



Extended Data Figure 8 | Example CSP. Temperature sensitivity (CSP_{temp}) for alderfly (*Sialis lutaria*) emergence from Windermere, UK. Solid black line: sensitivity of first emergence to water temperature on different days of the year (days change per $^{\circ}C$). Grey horizontal lines: 2.5th and 97.5th percentiles of these sensitivity values. Solid orange curve: GAM smoother

fitted through the sensitivity values with associated confidence intervals (dashed orange curves). Horizontal bars indicate where GAM confidence intervals exceed the percentiles of the original sensitivity values, indicating seasonal windows at the limits of the climate sensitivity profile ($n = 30$).

Extended Data Table 1 | Modelled relationships between seasonal timing and climate variables for $n=10,003$ phenological time series

Whole data set				Distribution of models among sensitivity types, for each taxonomic group										Distribution of models among sensitivity types, for each trophic level		
Climatic variable	CSP type	% series	% effects significant (lower, upper)	Amphibians	Birds	Crustacea	Fish	Insect	Mammals	Molluscs	Phytoplankton	Plants	Other	Prim. Prod.	Prim. Cons.	Sec. Cons.
Temperature	Type I	24.7	58.1, 0.4	27	21	19	45	26	27	20	12	20	24	16	26	21
	Type II	69.7	44.8, 1.0	73	68	68	55	69	60	60	76	77	65	76	69	68
	Type III	4.4	1.8, 2.2	0	8	10	0	4	12	20	10	3	8	7	4	7
Precipitation*	Type I	11.5	0.2, 0.1	27	16	0	18	11	5	NA†	11	13	10	11	11	16
	Type II	53.0	0.1, 0.6	55	50	67	73	52	66	NA†	67	58	70	63	52	51
	Type III	28.0	0.3, 0.6	9	22	14	0	30	22	NA†	18	23	17	21	30	22

CSPs fall within three broad response types; events always advance with increases in the climate variable irrespective of the seasonal timing of climate change (CSP Type I; Fig. 1, red curve), events are always delayed by increases in the climate variable irrespective of the seasonal timing of climate change (CSP Type III; Fig. 1, blue curve), and events may be advanced or delayed by increases in the climate variable, depending on the seasonal timing of climate change (CSP Type II; Fig. 1, orange curve). Shown are the percentage of series that fall in each type (% series), the percentage of effects that are statistically significant at $P<0.05$ after multiple testing correction (% effects significant). *Based only on freshwater and terrestrial taxa, for which precipitation data were available.

†NA indicates effect not evaluated, owing to lack of precipitation data for marine taxa.

Extended Data Table 2 | Parameter estimates and test statistics from climate–phenology mixed-effects models

	Response group	Intercept	Temp. Lower	Temp. Upper	Precip. Lower	Precip. Upper
Global Model	All species	169.1 (1.7, 98)	-2.6 (0.1, -35.2)	0.7 (0, 14.1)	-0.4 (0, -13)	1.4 (0.1, 19.4)
Trophic level model	Primary consumer (<i>b</i>)	211.6 (1.6, 131.1)	-3.7 (0.1, -39.5)	0.7 (0.1, 10.6)	-0.4 (0, -11.8)	2.2 (0.1, 23.7)
	Secondary consumer (Δb)	-67.5 (2.5, -26.5)	1.8 (0.1, 13.5)	0.0 (0.1, -0.2)	0.1 (0.1, 1.9)	-1.2 (0.1, -8.7)
	Primary producer (Δb)	-50.9 (5.7, -8.9)	-0.4 (0.3, -1.2)	0.7 (0.2, 3.3)	-0.3 (0.1, -2.4)	-0.4 (0.3, -1.2)
Taxonomic Class model	Phytoplankton (<i>b</i>)	199.2 (6.9, 28.9)	-4.1 (0.5, -7.8)	2.9 (0.2, 12.2)	-1.1 (0.1, -7.6)	2.5 (0.5, 5.4)
	Amphibians (Δb)	-115.9 (22.2, -5.2)	0.7 (1.2, 0.6)	-2.4 (1.2, -1.9)	0.9 (1, 0.9)	-2.1 (1.4, -1.5)
	Birds (Δb)	-58.5 (7.1, -8.2)	2.0 (0.5, 3.8)	-2.3 (0.3, -9)	0.7 (0.2, 4.6)	-1.4 (0.5, -3)
	Crustacea (Δb)	-17.1 (17.4, -1)	0.9 (0.9, 1)	-2.4 (0.9, -2.6)	0.7 (0.3, 2)	-1.0 (0.8, -1.2)
	Fishes (Δb)	-41.3 (23.8, -1.7)	1.1 (1.3, 0.9)	-2.4 (1.2, -2)	0.4 (0.6, 0.6)	-1.0 (1.2, -0.8)
	Insects (Δb)	16.9 (7, 2.4)	0.3 (0.5, 0.6)	-2.3 (0.2, -9.4)	0.6 (0.1, 4.3)	-0.3 (0.5, -0.6)
	Mammals (Δb)	43.8 (10.9, 4)	1.9 (0.7, 2.7)	-1.0 (0.5, -1.9)	0.8 (0.3, 2.6)	-1.2 (0.6, -1.8)
	Plants (Δb)	-64.0 (9.5, -6.8)	-0.2 (0.7, -0.3)	-3.5 (0.3, -11)	0.7 (0.2, 2.9)	-1.4 (0.7, -2)
Global Model*	All species	171.9 (1.9, 92.5)	-2.7 (0.1, -30.5)	0.7 (0.1, 10.2)		
Trophic level model*	Primary consumer (<i>b</i>)	214.4 (1.7, 124.3)	-3.8 (0.1, -31)	0.7 (0.1, 6.9)		
	Secondary consumer (Δb)	-66.6 (2.6, -25.3)	1.8 (0.2, 11.1)	0.0 (0.1, 0.2)		
	Primary producer (Δb)	-49.5 (5.2, -9.5)	-1.2 (0.4, -3.4)	1.8 (0.3, 5.8)		

Fixed-effect parameter estimates from each model are shown; the intercept and slope for each climatic predictor. Following *R* convention, absolute parameter estimates are provided for an assigned "baseline" group within each model (*b*), and remaining estimates are given as differences from this baseline (Δb). Each estimate has an associated standard error and *t* statistic in parentheses (standard error, *t*). Climatic predictors include mean temperature and precipitation in seasonal windows at the upper and lower limit of the climate sensitivity profile for each phenological series. The number of observations $n = 370,725$. *Models were re-run including the marine plankton data, and excluding precipitation effects (see text). In these models the number of observations $n = 379,081$.

Multiple mechanisms disrupt the *let-7* microRNA family in neuroblastoma

John T. Powers¹, Kaloyan M. Tsanov¹, Daniel S. Pearson¹, Frederik Roels², Catherine S. Spina³, Richard Ebricht¹, Marc Seligson¹, Yvanka de Soysa¹, Patrick Cahan¹, Jessica Theißen², Ho-Chou Tu¹, Areum Han¹, Kyle C. Kurek⁴, Grace S. LaPier¹, Jihan K. Osborne¹, Samantha J. Ross¹, Marcella Cesana¹, James J. Collins^{3,5}, Frank Berthold² & George Q. Daley^{1,6,7,8}

Poor prognosis in neuroblastoma is associated with genetic amplification of *MYCN*. *MYCN* is itself a target of *let-7*, a tumour suppressor family of microRNAs implicated in numerous cancers. *LIN28B*, an inhibitor of *let-7* biogenesis, is overexpressed in neuroblastoma and has been reported to regulate *MYCN*. Here we show, however, that *LIN28B* is dispensable in *MYCN*-amplified neuroblastoma cell lines, despite de-repression of *let-7*. We further demonstrate that *MYCN* messenger RNA levels in amplified disease are exceptionally high and sufficient to sponge *let-7*, which reconciles the dispensability of *LIN28B*. We found that genetic loss of *let-7* is common in neuroblastoma, inversely associated with *MYCN* amplification, and independently associated with poor outcomes, providing a rationale for chromosomal loss patterns in neuroblastoma. We propose that *let-7* disruption by *LIN28B*, *MYCN* sponging, or genetic loss is a unifying mechanism of neuroblastoma development with broad implications for cancer pathogenesis.

Carcinogenesis involves multiple genetic and epigenetic events, yet the organizing principles underlying their choreography are poorly understood. MicroRNA (miRNA) deregulation is an important component of this landscape through both oncogenic and tumour-suppressive functions of miRNAs¹. Of these, the highly conserved *let-7* family has a prominent role in the regulation of embryonic development and maintenance of differentiated tissues and is among the most abundantly expressed miRNAs. It serves as a potent tumour suppressor via post-transcriptional repression of multiple oncogenic messenger RNA (mRNA) targets including *RAS*, *MYC*, and *HMG2* (refs 2–4). The *let-7* family is downregulated in multiple tumour types and has been causally linked to oncogenesis^{1,5–8}. Uncovering the mechanisms by which *let-7* function is neutralized is therefore critical to both the fundamental understanding of cancer pathogenesis and novel therapies.

Several mechanisms of *let-7* disruption have emerged in different contexts. First, its biogenesis can be suppressed by the *LIN28B* RNA-binding protein⁹, a highly conserved heterochronic gene implicated in cancer and reported to induce tumours in multiple mouse models including hepatocellular carcinoma, colon cancer, Wilms tumour, and neuroblastoma^{10–16}. Second, competing endogenous RNAs (ceRNAs) have been proposed to sponge miRNAs, including *let-7*, diluting their activity through competition for miRNAs with sites common to multiple ceRNA species^{17–19}. Third, chromosome loss is a suggested mechanism of *let-7* disruption in cancer, as genetic deletion of *let-7* is associated with several solid tumours¹.

The neuroblastoma master oncogene, *MYCN*, has a 910-nucleotide-long 3' untranslated region (UTR) containing two *let-7* binding sites which are almost perfectly conserved among land vertebrates, suggesting strong functional relevance^{20–22} (Extended Data Fig. 1). Coding sequence mutations in neuroblastoma are rare^{23,24}, whereas chromosome arm gain or loss events are common^{25,26}. The most well-known chromosomal aberration is amplification of the *MYCN* locus, which occurs in ~25% of all neuroblastomas and largely defines poor

prognosis^{27,28}. Other common chromosomal deletions at chromosome arms 3p and 11q are inversely associated with *MYCN* amplification. The reason for this discordance is unknown.

Here we set out to understand the relationship between *MYCN* and *let-7* in neuroblastoma. A complex relationship emerges between *LIN28B* activity, a novel ceRNA function of the *MYCN* 3' UTR, and *let-7* genetic loss, which together present a unifying model of *let-7* suppression during neuroblastoma pathogenesis. This model provides an organizing principle for understanding distinct genetic patterning in neuroblastoma, with potential implications for cancer in general.

LIN28B and *let-7* regulate the *MYCN* 3' UTR

LIN28B is highly expressed in human neuroblastoma and its expression correlates with tumour stage, rendering the *LIN28B/let-7* axis an attractive target for interrogation (Extended Data Fig. 2a–d). Two recent reports concluded that this pathway plays a critical role in regulating *MYCN* and neuroblastoma cell growth^{12,13}. To examine the relationship between the *MYCN* transcript, *let-7* and *LIN28B*, we first transfected non-*MYCN* amplified neuroblastoma cells with the *MYCN* open reading frame (ORF), with or without the 3' UTR carrying intact or mutant *let-7* sites (Fig. 1a). The full-length wild-type *MYCN* transcript produced markedly lower *MYCN* protein levels than the ORF-only construct. Mutation of the *let-7* sites in the 3' UTR partly rescued *MYCN* expression, implicating *let-7* modulation as an important component of *MYCN* post-transcriptional regulation (Fig. 1b). Expression of *LIN28B* suppressed the *let-7* family in non-*MYCN*-amplified neuroblastoma cells and conferred a growth advantage. *LIN28B* rescued expression of the wild-type 3' UTR construct, demonstrating that *LIN28B* can support *MYCN* expression through *let-7* repression in the absence of *MYCN* amplification (Extended Data Fig. 2e, f and Fig. 1c). However, when we transfected *MYCN*-amplified cells with a *let-7a* mimic, we observed decreased *MYCN* protein levels only above 15- and 80-fold increases in cellular levels of *let-7a*, respectively,

¹Division of Pediatric Hematology/Oncology, Boston Children's Hospital, Boston, Massachusetts 02115, USA. ²Department of Pediatric Oncology, University Hospital Köln, Köln 50937, Germany.

³Wyss Institute for Biologically Inspired Engineering, Boston, Massachusetts 02115, USA. ⁴Department of Pathology, Boston Children's Hospital, Boston Massachusetts 02115, USA. ⁵Department of Biological Engineering, Massachusetts Institute of Technology, Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. ⁶Stem Cell Transplantation Program, Dana Farber Cancer Institute & Boston Children's Hospital, Boston, Massachusetts 02115, USA. ⁷Department of Biological Chemistry & Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁸Harvard Stem Cell Institute, Cambridge, Massachusetts 02138, USA.

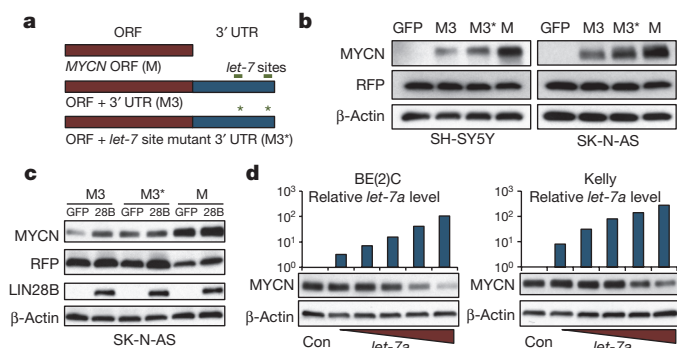


Figure 1 | The *LIN28B*/*let-7*/*MYCN* axis is intact in neuroblastoma. **a**, pcDNA3.1:*MYCN* constructs. **b**, Immunoblot for indicated proteins in non-*MYCN*-amplified neuroblastoma cells transfected with indicated constructs. **c**, Immunoblot for indicated proteins in green fluorescent protein (*GFP*)- or *LIN28B*-expressing cells transfected as in **b**. **d**, qPCR analysis of *let-7a* levels and immunoblot for *MYCN* in *MYCN*-amplified cells transfected with control or *let-7a* mimic dose curve (data representative of three independent experiments). For gel source data, see Supplementary Figures.

suggesting that *MYCN* was refractory to all but exceedingly high levels of exogenous *let-7* (Fig. 1d).

LIN28B is dispensable in *MYCN*-amplified cells

Next, we evaluated the previously reported *LIN28B*-*let-7*-*MYCN* regulatory circuit using published lentiviral short hairpin RNA (shRNA) constructs to knockdown *LIN28B* in *MYCN*-amplified neuroblastoma cells and observed comparable suppression of *MYCN* protein levels and cell growth (Extended Data Fig. 3a, b). We further observed reduced xenograft tumour growth in cells expressing a *LIN28B* targeting shRNA (Extended Data Fig. 3c). However, we did not observe an appreciable de-repression of *let-7* levels upon shRNA-mediated *LIN28B* knockdown, which is counter to the established model (Extended Data Fig. 3d). Moreover, we were unable to rescue these effects through overexpression of shRNA-resistant *LIN28B* constructs (Extended Data Fig. 3e, f). Together, these data suggest that the reported effects of the shRNAs on both cell growth and *MYCN* protein levels might be due to hairpin-induced toxicities.

As an alternative approach to depleting *LIN28B*, we tested five small interfering RNAs (siRNAs) and found that four both effectively knocked down *LIN28B* and, as expected, de-repressed *let-7* levels (Extended Data Fig. 4a–d). Upon extended serial siRNA transfection, we observed that despite robust *LIN28B* knockdown and strong de-repression of *let-7*, *MYCN* protein levels were unaffected and there was no appreciable effect on cell proliferation (Extended Data Fig. 4e–g).

To exclude the possibility of incomplete knockdown resulting in residual *LIN28B* activity, we employed *Cas9* and four distinct guide RNAs (gRNAs) targeting *LIN28B* (Extended Data Fig. 4h). We observed robust loss of *LIN28B* protein with all four gRNA constructs (Fig. 2a, b), indicating efficient disruption of the locus. We did not observe appreciable loss of *MYCN* protein expression or impaired cell growth, thus corroborating our siRNA-based results (Fig. 2a–d). In addition, the *let-7* family was robustly de-repressed, consistent with the existing *LIN28B*/*let-7* model (Fig. 2e, f). These observations indicate that disruption of *LIN28B* has little net impact on *MYCN*-amplified neuroblastoma cells.

Amplified *MYCN* mRNA is a *let-7* sponge

The persistence of *MYCN* protein levels despite high levels of transfected *let-7a* or robust de-repression of *let-7* upon *LIN28B* loss prompted us to explore novel mechanisms of *let-7* perturbation. We hypothesized that there might be a ceRNA in *MYCN*-amplified neuroblastoma cells that serves to sponge *let-7*. To identify potential ceRNAs, we performed poly-A selected RNA-sequencing (mRNA-seq)

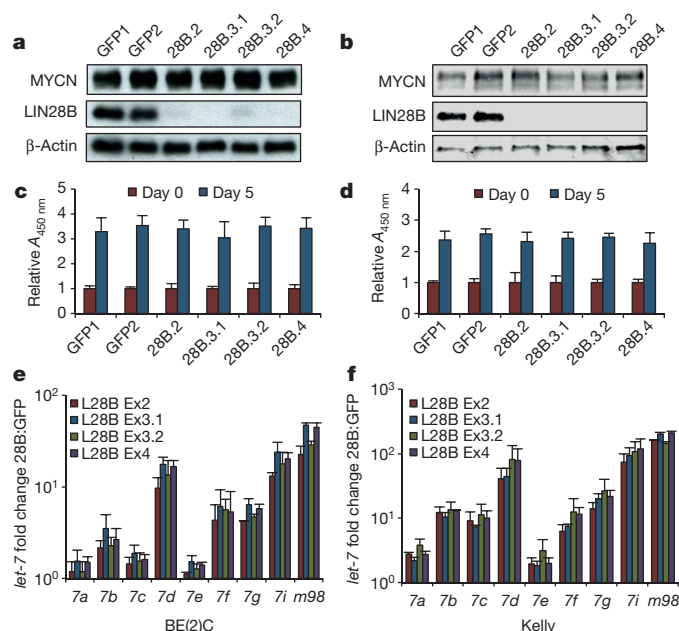


Figure 2 | *LIN28B* is dispensable in human *MYCN* amplified neuroblastoma cells. **a**, **b**, Immunoblot for *MYCN* and *LIN28B* in cells infected with indicated *Cas9*-gRNA lentivirus. For gel source data, see Supplementary Figures. **c**, **d**, Cell growth analysis of cells infected as in **a**, **b** (BrdU incorporation; relative absorbance, *A*, at 450 nm). **e**, **f**, qPCR analysis of relative *let-7* expression in cells from **a** and **b** respectively. Fold change relative to *GFP* gRNA controls (*n* = 3 independent experiments, mean + s.e.m. shown).

on *MYCN*-amplified (BE(2)C and Kelly) and non-*MYCN*-amplified (SH-SY5Y) cells. We then determined the relative contribution of *let-7* target sites provided by expressed *let-7* targets. Interestingly, *MYCN* itself was by far the most abundant *let-7* target mRNA in both BE(2)C and Kelly cells, alone providing 19.3% and 18.5% of the entire cellular *let-7* target-site pool, respectively. In contrast, *MYCN* represented only 0.15% of the *let-7* target-site pool in SH-SY5Y cells (Fig. 3a). In fact, *MYCN* mRNA was the second highest expressed mRNA in both BE(2)C and Kelly cells as opposed to the 5,409th highest in SH-SY5Y, demonstrating an exceptionally high *MYCN* mRNA level in *MYCN*-amplified cells (>100-fold increase; Fig. 3b). Other multiple-*let-7*-site mRNAs such as *HMG2*, *IMPI*, and *ARID3B* were expressed at much lower levels, together suggesting that *MYCN* mRNA might itself be the sponge (Extended Data Fig. 5a). This expression pattern was validated by quantitative PCR (qPCR) in a panel of additional cell lines (Extended Data Fig. 5b).

ceRNA relationships were initially defined in part by similar expression levels between RNAs with similar 3' UTRs^{18,29}. Two recent reports have refined this original precept, suggesting that for a given miRNA family, the miRNA:mRNA-target ratio is a major determinant of how effectively a ceRNA can impact the function of a miRNA family. At low ratios, miRNAs are sensitive to moderate levels of ceRNAs, whereas highly expressed miRNAs with high ratios are difficult to sponge, requiring very high levels of ceRNA^{30,31}. We therefore assessed total copies per cell of both *MYCN* mRNA and the *let-7* family in BE(2)C and Kelly cells through quantified mRNA-seq and small-RNA sequencing (sRNA-seq) (Extended Data Fig. 6).

We calculated 13,255 and 10,615 *MYCN* mRNA copies per BE(2)C and Kelly cell, respectively, resulting in 26,511 and 21,231 *let-7* target sites provided by *MYCN* (Fig. 3c top). In stark contrast, there were only 31 *MYCN* transcripts per SH-SY5Y cell. Quantification of the *let-7* family in BE(2)C and Kelly cells yielded 7,259 and 1,952 total *let-7* molecules per cell (Fig. 3c bottom), and *MYCN*-*let-7*-site:*let-7* miRNA ratios of 3.65 and 10.88, respectively. These ratios satisfy the tenets of the ratio-based ceRNA model and do so as the result of a single

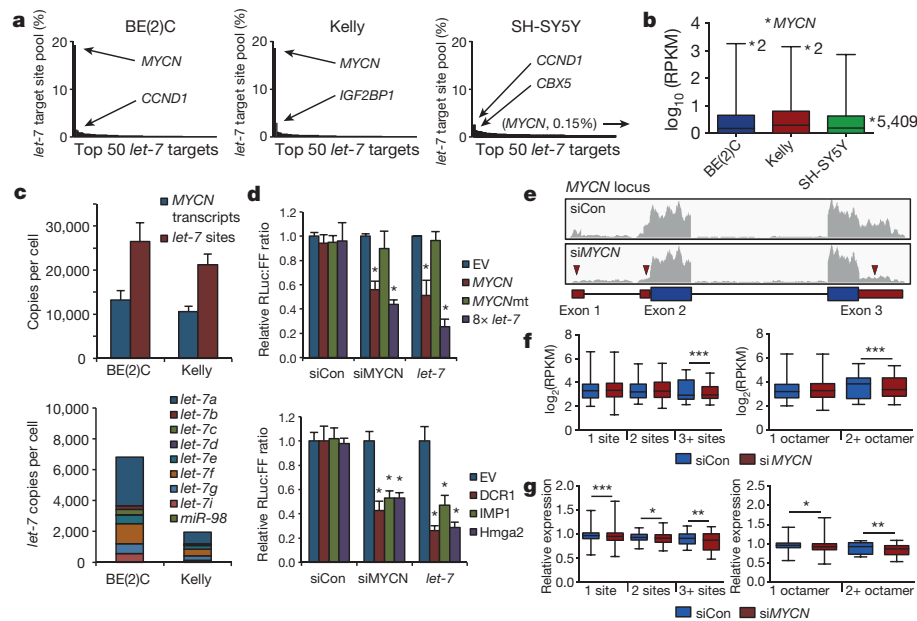


Figure 3 | MYCN mRNA is a ceRNA for *let-7*. **a**, The *let-7* target-site pool contribution of the top 50 targets. Arrows indicate top two contributors ($n = 3$ biological replicates, Gene Expression Omnibus accession number GSE81498, see Source Data F3 in Supplementary Information). **b**, Box and whisker plot of top 14,000 expressed mRNAs from data set described in **a**. MYCN expression rank marked by asterisk (GSE81498, see Source Data F3). **c**, Top: MYCN transcripts and *let-7* sites provided per cell as determined by the data set from **a** (mean \pm s.d. shown, GSE81498, see Source Data F3). Bottom: *let-7* copies per cell as determined by sRNA-seq ($n = 3$ biological replicates, GSE81499, see Source Data F3). **d**, Top: relative fluorescence ratio in BE(2)C cells co-transfected with indicated constructs small RNAs. Bottom: relative fluorescence ratio in cells co-transfected with indicated constructs and small RNAs. Mean of four independent experiments plus s.e.m.

amplified mRNA. Interestingly, *let-7a* was the most highly expressed *let-7* family member in both cell types, accounting for over half of all *let-7* molecules. These observations were confirmed by spike-in qPCR-based quantification of both MYCN and *let-7* (Extended Data Fig. 7a, b). Even upon LIN28B knockout, the MYCN-*let-7*-site:*let-7* ratio is 1.35 in BE(2)C and 1.78 in Kelly, which remain favourable for ceRNA activity (Extended Data Fig. 7c).

To test the capacity of MYCN mRNA to serve as a *let-7* sponge, we co-transfected BE(2)C cells with a series of luciferase constructs containing the 3' UTRs of several representative *let-7* targets and control or MYCN ORF targeting siRNA (Extended Data Fig. 8a, b). Luciferase ratios of all constructs except for empty vector controls and the *let-7*-site-mutated MYCN-3' UTR were significantly reduced by either MYCN knockdown or *let-7a* transfection (Fig. 3d), suggesting that the endogenous MYCN-3' UTR sponges steady-state levels of *let-7*. We then tested the sufficiency of the MYCN-3' UTR to de-repress *let-7* targets through sponging of a *let-7a* mimic. We co-transfected the above luciferase reporter constructs with chimaeric RFP:MYCN-3' UTR constructs and assayed luciferase activity (Extended Data Fig. 8c, d). The *let-7* target constructs were rescued when co-transfected with wild-type but not *let-7*-site-mutant MYCN-3' UTR (Extended Data Fig. 8e). In addition, exogenous MYCN-3' UTR was sufficient to enhance MYCN protein expression itself in SK-N-AS cells (Extended Data Fig. 8f).

We next tested whether endogenous *let-7* targets are similarly affected. Upon MYCN knockdown, protein levels of DICER1, HK2, IMP1, and LIN28B were reduced, while neither mRNA nor *let-7* levels were significantly changed (Extended Data Fig. 9a–c). Concurrent *let-7* inhibition rescued expression of the four targets, supporting post-transcriptional suppression through *let-7* upon MYCN knockdown that

($*P < 0.05$ relative to empty vector, unpaired *t*-test). **e**, mRNA-seq reads mapping to the MYCN locus in BE(2)C:MYCN cells transfected with indicated siRNAs. Blue and red boxes indicate MYCN ORFs and UTRs, respectively. Triangles mark reduction of MYCN mRNA UTRs. **f**, Expression levels of mRNAs with one, two, or three or more *let-7* sites (left) or with one or two or more octamer *let-7* sites (right) in siCon and siMYCN cells (data represent one round of mRNA-seq, $***P < 0.0001$, one-tailed Wilcoxon test). **g**, Relative expression of mRNAs with one, two, or three or more *let-7* sites (left) or with one or two or more octamer *let-7* sites (right) in siCon and siMYCN cells co-transfected with *let-7a*. Values relative to miRCon. (Data represent one round of mRNA-seq, $*P < 0.05$, $**P < 0.001$, $***P < 0.001$ versus siCon, one-tailed Wilcoxon test, GSE81497, see Source Data F3).

is consistent with MYCN mRNA serving as a *let-7* sponge (Extended Data Fig. 9a). Further, these targets were not reduced at the protein level upon LIN28B knockout, which is consistent with this model (Extended Data Fig. 9d).

Lastly, we analysed global *let-7* target expression in response to depletion of the endogenous MYCN 3' UTR. To specifically assess the role of the 3' UTR, we transfected BE(2)C cells expressing a MYCN ORF transgene with a MYCN 3' UTR-targeting siRNA (Fig. 3e). Although we did not see a global reduction of *let-7* targets as a whole (Extended Data Fig. 9e), we did observe significantly lower expression of *let-7* targets with three or more total sites or more than one octamer *let-7* site in their 3' UTRs, which together define the most sensitive *let-7* targets^{32,33} (Fig. 3f). Given the regulation of the MYCN 3' UTR by multiple miRNAs³⁴ and unknown kinetics of how *let-7* activity is restored after the removal of a ceRNA, we further challenged siCon and siMYCN cells with a modest amount of *let-7a* mimic, increasing cellular *let-7* levels approximately eightfold (Extended Data Fig. 9f). We then observed significantly reduced expression across all categories of *let-7* targets in siMYCN cells, consistent with increased sensitivity to *let-7* in the absence of the MYCN 3' UTR (Fig. 3g and Extended Data Fig. 9g).

Chromosomal loss of *let-7* in neuroblastoma

While neuroblastoma has a low mutation rate, chromosome arm gain and loss is frequent^{23,24}. Two of the most common chromosomal losses in neuroblastoma, chromosome arm 3p (Chr3p, ~33% incidence) and Chr11q (~45% incidence), often occur together and seldom with MYCN amplification^{23,35,36} (Fig. 4a and Extended Data Fig. 10). Upon analysis of Chr3p and Chr11q, we noted that the Chr3p-loss smallest region of overlap spans from 3p25.3 to 3p14.3 (ref. 37), placing *let-7g*

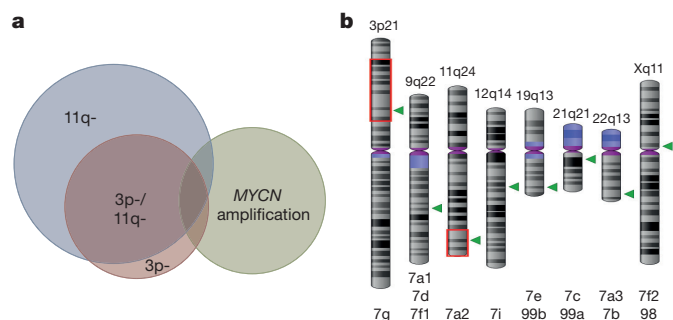


Figure 4 | Loss of *let-7a2* and *let-7g* is common in neuroblastoma. **a**, Scaled Venn diagram assembled from ref. 23 detailing relative incidence and intersection of chromosome 11q and 3p loss and MYCN amplification. **b**, Genetic locations of 12 distinct *let-7* family members. Green triangles, *let-7* loci. Red boxes indicate the smallest region of overlap for these deletions. (Images created at <http://www.ncbi.nlm.nih.gov/genome/tools/gdp/>.)

within the smallest region of overlap and resulting in its loss whenever Chr3p is lost. In addition, the most common breakpoint of Chr11q immediately proximal to the *let-7a2* locus, resulting in loss of *let-7a2* in virtually all Chr11q deletions³⁵ (Fig. 4b). Moreover, loss of Chr11q in neuroblastoma results in lower overall *let-7a* levels despite loss of only one of three distinct *let-7a* loci³⁸. These observations suggest that *let-7* genetic disruption may be selected for in neuroblastoma.

A model in which amplified MYCN sequesters mature *let-7* would predict that selective pressure to genetically lose *let-7* might be relieved in MYCN-amplified disease. Chr3p (*let-7g*) and Chr11q (*let-7a2*) loss patterns are indeed consistent with such a model. To investigate whether the extended *let-7* family (Fig. 4b) follows this loss pattern, we expanded our analysis to all eight *let-7* genetic loci in 202 neuroblastomas by array comparative genomic hybridization (aCGH). We created a heat-map of copy number estimates for each miRNA locus to compare MYCN-amplified with non-amplified copy number values and observed a significant difference for six *let-7* loci (Fig. 5a, b top). At least one *let-7* family member was lost in 63.4% of non-amplified tumours and in only 16.7% of amplified (Fig. 5b top), resulting in average *let-7* copy number changes of -1.94 and -0.36 per tumour, respectively (Fig. 5b bottom). This pattern of copy number loss for *let-7* is distinct from the unrelated *miR-103a* family (Fig. 5b), suggesting that the *let-7* pattern is not reflective of general chromosomal patterning.

The most commonly lost were *let-7a2*, *let-7f2*, and *let-7g*, whereas *let-7a3/7b* and *let-7i* were not significantly lost in any tumour subset. We reasoned that loss frequency might relate to initial expression level. To interrogate this possibility, we used publicly available sRNA-seq data to examine the relative expression levels of mature *let-7* in 12 distinct primary and tumour cell lines that have intact *let-7* loci (Extended Data Fig. 10b). We observed that *let-7a*, *let-7f*, and *let-7g* are present at higher relative levels than other *let-7* family members, which mirrors the copy number loss patterns in non-MYCN-amplified neuroblastoma (Fig. 5c). *MIR-100*, the miRNA-cluster partner of *let-7a2*, is more highly expressed than most *let-7s* and their cluster partners, suggesting that

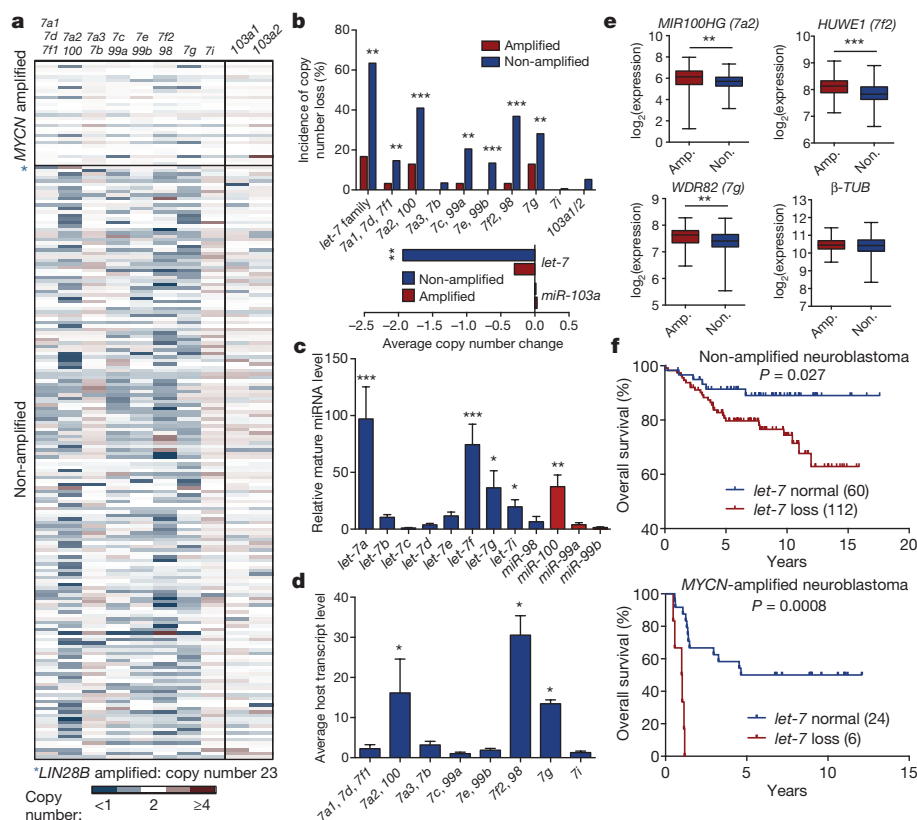


Figure 5 | The MYCN acRNA model predicts *let-7* chromosomal loss patterns in neuroblastoma. **a**, The *let-7* and *miR-103a* loci copy number heat-map based on aCGH relative fluorescence ratios of tumour versus germline (Source Data F5 in Supplementary Information). **b**, Copy number loss incidence for *let-7* and *miR-103a* from data set in **a** (top). Average copy number change for the two miRNA families (bottom). (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ amplified versus non-amplified for each locus, unpaired *t*-test, Source Data F5.) **c**, Mature *let-7* expression based on ENCODE sRNA-seq data. (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$

versus *let-7c*, Wilcoxon test, Source Data F5.) **d**, Relative *let-7* host transcript levels based on ENCODE mRNA-seq data. (* $P < 0.05$, ** $P < 0.01$ versus *let-7c*, Wilcoxon test, Source Data F3.) **e**, Relative expression of indicated host genes in MYCN-amplified (Amp.) and non-amplified (Non.) neuroblastoma. β -TUB shown as control ($n = 498$, ** $P < 0.01$, *** $P < 0.001$, unpaired *t*-test, Source Data F5). **f**, Overall survival curves of neuroblastoma patients in non-amplified (top), and MYCN-amplified (bottom) neuroblastoma. *P* values determined by Mantel-Cox test (Source Data F5).

the bulk of the *let-7a* reads may come from the *let-7a2/miR-100* locus (Fig. 5c).

A limitation of sRNA-seq is that it cannot distinguish between loci that produce the same mature miRNA. Family members of *let-7* are coordinately transcribed as part of larger host transcripts from which they are then processed, which allows for locus-specific expression analysis^{7,39,40} (Extended Data Fig. 10c). We therefore analysed relative expression levels of *let-7* host transcripts in six primary and tumour cell lines. Host transcript levels for *let-7a2*, *let-7f2*, and *let-7g* were significantly higher than for other *let-7* host transcripts (Extended Data Fig. 10d and Fig. 5d), reflecting the pattern of most frequent locus loss in non-amplified disease. In addition, analysis of existing human neuroblastoma mRNA-seq and microarray data sets revealed lower expression of the *let-7a2*, *let-7f2*, and *let-7g* host transcripts in non-amplified compared with *MYCN*-amplified tumours, which is consistent with observed patterns of copy number loss in our aCGH data set (Fig. 5e and Extended Data Fig. 10e). Further, *MYCN* and *let-7* expression are negatively correlated in non-amplified disease, underscoring the importance of *let-7* disruption in the absence of the *MYCN* ceRNA³⁴. These data may collectively explain both preferential loss of certain *let-7* loci and common patterns of chromosomal loss in neuroblastoma.

Further emphasizing the significance of *let-7* suppression is the observation that non-*MYCN*-amplified neuroblastoma patients had significantly worse overall survival if there was a *let-7* copy number loss event (Fig. 5f top). In the rare case where *MYCN* amplification and *let-7* copy number loss occurred together, overall survival was dramatically reduced relative to the already-poor prognosis of *MYCN* amplification (Fig. 5f bottom), suggesting a deleterious synergy between two powerful but typically exclusive mechanisms of functional *let-7* disruption.

Of note, one of the tumours in the aCGH data set had genetic amplification of *LIN28B* (copy number = 23; Fig. 5a). This tumour patterned closely with *MYCN*-amplified tumours with regard to net *let-7* loss and tumour stage (IV) despite being 2n for *MYCN*. This may represent a case where *LIN28B* significantly contributed to neuroblastoma through *let-7* suppression, similar to the reported mouse model of murine-*Lin28b* driven neuroblastoma¹².

Discussion

The known functionality of *LIN28B*, together with the patterns of genetic deletion of *let-7* and amplified *MYCN* ceRNA (aceRNA) activity described here, establish that neuroblastoma employs multiple mechanisms to neutralize *let-7*, placing *let-7* disruption at the centre of neuroblastoma pathogenesis. We thus propose that *let-7* biogenesis and function are targeted in neuroblastoma by several disparate mechanisms: high frequency genetic loss, *LIN28B* activity, or *MYCN* aceRNA (Extended Data Fig. 10f). This model has implications for our understanding of neuroblastoma pathogenesis, disease modelling, and the rational design of therapeutic strategies, and may represent a more general feature of human cancer.

First, our model offers a plausible explanation for the uniquely high *MYCN* mRNA levels in amplified neuroblastoma, which enable both robust expression of *MYCN* protein and adequate copies of a ceRNA sufficient to impair the function of a highly expressed miRNA such as *let-7*. Questions may remain about whether an observed >100-fold increase in *MYCN* mRNA, which increases the total *let-7* target sites across the cellular pool of mRNAs by only approximately 25%, is sufficient to mediate a ceRNA effect on *let-7*. However, our functional data based on loss of the *MYCN* 3' UTR, including candidate reporter analysis of *let-7*-site-containing 3' UTRs and global *let-7*-target mRNA-seq analysis (Fig. 3d–g), suggest that *MYCN* mRNA may be a preferred target that in abundance can sequester and impair *let-7*. Further, in tumours lacking *MYCN* amplification, our model suggests that selective pressure to disrupt *let-7* explains the well-known, yet unresolved, patterns of *MYCN*-amplification-independent chromosome 3p and 11q loss. AceRNA function of *MYCN* mRNA also accounts for the

dispensability of *LIN28B* in *MYCN*-amplified cell lines, suggesting that *LIN28B* may serve a redundant *let-7* suppressive role.

Our findings suggest that highly expressed 3' UTRs contribute to miRNA deregulation in cancer, and therefore both coding and non-coding functions of oncogenic mRNAs should be considered in animal tumour modelling. For example, both the TH-*MYCN* and LSL-*MYCN*; *Dbh-iCre* models of murine neuroblastoma overexpress the *MYCN* ORF and lack 3' UTRs. Notably, the TH-*MYCN* model has similar patterns of *let-7a2*, *let-7f2*, and *let-7g* genetic loss as non-amplified human disease and both models broadly downregulate the *let-7* family^{34,41}, further suggesting that *let-7* disruption is important even in the presence of *MYCN* protein expression. Moreover, a recent report demonstrates that high-level expression of the *BRAF* pseudogene, which contains a functional 3' UTR but does not encode a protein, is sufficient to induce lymphoma in mice⁴². Consequently, full-length mRNA transgenes may yield more accurate genetic modelling of human tumours in animals.

Lastly, our model establishes *let-7* restoration as a key therapeutic goal in neuroblastoma. There are few neuroblastoma-specific therapies, and attempts to directly target *MYCN* have met with little success, despite efforts spanning the past 20 years (refs 25, 26, 43). The fact that *MYCN* mRNA has such a strong functional connection to *let-7* exposes a valuable opportunity to target *MYCN* itself and provides hope of delivering disease-specific therapy to the worst prognostic class of neuroblastoma.

We show here that disparate modes of *let-7* suppression are selectively and inversely related in neuroblastoma. Given that both oncogenic amplification and disruption of *let-7* biogenesis appear to play central roles in multiple cancer types^{15,44–47}, our model may provide a novel organizing principle by which to consider and interrogate genetic events in a broad range of tumours.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 2 February 2015; accepted 8 June 2016.

Published online 6 July 2016.

1. Croce, C. M. Causes and consequences of microRNA dysregulation in cancer. *Nature Rev. Genet.* **10**, 704–714 (2009).
2. Johnson, S. M. et al. RAS is regulated by the let-7 microRNA family. *Cell* **120**, 635–647 (2005).
3. Sampson, V. B. et al. MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. *Cancer Res.* **67**, 9762–9770 (2007).
4. Mayr, C., Hemann, M. T. & Bartel, D. P. Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science* **315**, 1576–1579 (2007).
5. Lu, J. et al. MicroRNA expression profiles classify human cancers. *Nature* **435**, 834–838 (2005).
6. Boyerinas, B., Park, S.-M., Hau, A., Murmann, A. E. & Peter, M. E. The role of let-7 in cell differentiation and cancer. *Endocr. Relat. Cancer* **17**, F19–F36 (2010).
7. Gurtan, A. M. & Sharp, P. A. The role of miRNAs in regulating gene expression networks. *J. Mol. Biol.* **425**, 3582–3600 (2013).
8. Blandino, G. et al. Tumor suppressor microRNAs: a novel non-coding alliance against cancer. *FEBS Lett.* **588**, 2639–2652 (2014).
9. Viswanathan, S. R., Daley, G. Q. & Gregory, R. I. Selective blockade of microRNA processing by Lin28. *Science* **320**, 97–100 (2008).
10. Nguyen, L. H. et al. Lin28b is sufficient to drive liver cancer and necessary for its maintenance in murine models. *Cancer Cell* **26**, 248–261 (2014).
11. Viswanathan, S. R. et al. Lin28 promotes transformation and is associated with advanced human malignancies. *Nature Genet.* **41**, 843–848 (2009).
12. Molenaar, J. J. et al. LIN28B induces neuroblastoma and enhances MYCN levels via let-7 suppression. *Nature Genet.* **44**, 1199–1206 (2012).
13. Diskin, S. J. et al. Common variation at 6q16 within HACE1 and LIN28B influences susceptibility to neuroblastoma. *Nature Genet.* **44**, 1126–1130 (2012).
14. Madison, B. B. et al. LIN28B promotes growth and tumorigenesis of the intestinal epithelium via Let-7. *Genes Dev.* **27**, 2233–2245 (2013).
15. Urbach, A. et al. Lin28 sustains early renal progenitors and induces Wilms tumor. *Genes Dev.* **28**, 971–982 (2014).
16. Tu, H. C. et al. LIN28 cooperates with WNT signaling to drive invasive intestinal and colorectal adenocarcinoma in mice and humans. *Genes Dev.* **29**, 1074–1086 (2015).

17. Tay, Y., Rinn, J. & Pandolfi, P. P. The multilayered complexity of ceRNA crosstalk and competition. *Nature* **505**, 344–352 (2014).
18. Poliseno, L. *et al.* A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**, 1033–1038 (2010).
19. Cesana, M. & Daley, G. Q. Deciphering the rules of ceRNA networks. *Proc. Natl Acad. Sci. USA* **110**, 7112–7113 (2013).
20. Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20 (2005).
21. Melton, C., Judson, R. L. & Blelloch, R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature* **463**, 621–626 (2010).
22. Baeyens, K. J., De Bondt, H. L., Pardi, A. & Holbrook, S. R. A curved RNA helix incorporating an internal loop with G-A and A-A non-Watson-Crick base pairing. *Proc. Natl Acad. Sci. USA* **93**, 12851–12855 (1996).
23. Pugh, T. J. *et al.* The genetic landscape of high-risk neuroblastoma. *Nature Genet.* **45**, 279–284 (2013).
24. Molenaar, J. J. *et al.* Sequencing of neuroblastoma identifies chromothripsis and defects in neurogenesis genes. *Nature* **483**, 589–593 (2012).
25. Barone, G., Anderson, J., Pearson, A. D. J., Petrie, K. & Chesler, L. New strategies in neuroblastoma: therapeutic targeting of MYCN and ALK. *Clin. Cancer Res.* **19**, 5814–5421 (2013).
26. Maris, J. M. Recent advances in neuroblastoma. *N. Engl. J. Med.* **362**, 2202–2211 (2010).
27. Seeger, R. C. *et al.* Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. *N. Engl. J. Med.* **313**, 1111–1116 (1985).
28. Brodeur, G. M., Seeger, R. C., Schwab, M., Varmus, H. E. & Bishop, J. M. Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. *Science* **224**, 1121–1124 (1984).
29. Ala, U. *et al.* Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments. *Proc. Natl Acad. Sci. USA* **110**, 7154–7159 (2013).
30. Denzler, R., Agarwal, V., Stefano, J., Bartel, D. P. & Stoffel, M. Assessing the ceRNA hypothesis with quantitative measurements of miRNA and target abundance. *Mol. Cell* **54**, 766–776 (2014).
31. Bosson, A. D., Zamudio, J. R. & Sharp, P. A. Endogenous miRNA and target concentrations determine susceptibility to potential ceRNA competition. *Mol. Cell* **56**, 347–359 (2014).
32. Agarwal, V., Bell, G. W., Nam, J.-W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **4**, <http://dx.doi.org/10.7554/eLife.05005> (2015).
33. Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
34. Beckers, A. *et al.* MYCN-targeting miRNAs are predominantly downregulated during MYCN-driven neuroblastoma tumor formation. *Oncotarget* **6**, 5204–5216 (2015).
35. Maris, J. M. *et al.* Allelic deletion at chromosome bands 11q14-23 is common in neuroblastoma. *Med. Pediatr. Oncol.* **36**, 24–27 (2001).
36. Breen, C. J., O'Meara, A., McDermott, M., Mullarkey, M. & Stallings, R. L. Coordinate deletion of chromosome 3p and 11q in neuroblastoma detected by comparative genomic hybridization. *Cancer Genet. Cytogenet.* **120**, 44–49 (2000).
37. Ejeskär, K., Aburatani, H., Abrahamsson, J., Kogner, P. & Martinsson, T. Loss of heterozygosity of 3p markers in neuroblastoma tumours implicate a tumour-suppressor locus distal to the FHIT gene. *Br. J. Cancer* **77**, 1787–1791 (1998).
38. Bray, I. *et al.* Widespread dysregulation of MiRNAs by MYCN amplification and chromosomal imbalances in neuroblastoma: association of miRNA expression with survival. *PLoS ONE* **4**, e7850 (2009).
39. Chiang, H. R. *et al.* Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev.* **24**, 992–1009 (2010).
40. Roush, S. & Slack, F. J. The let-7 family of microRNAs. *Trends Cell Biol.* **18**, 505–516 (2008).
41. Hackett, C. S. *et al.* Genome-wide array CGH analysis of murine neuroblastoma reveals distinct genomic aberrations which parallel those in human tumors. *Cancer Res.* **63**, 5266–5273 (2003).
42. Karreth, F. A. *et al.* The BRAF pseudogene functions as a competitive endogenous RNA and induces lymphoma *in vivo*. *Cell* **161**, 319–332 (2015).
43. Maris, J. M., Hogarty, M. D., Bagatell, R. & Cohn, S. L. Neuroblastoma. *Lancet* **369**, 2106–2120 (2007).
44. Pugh, T. J. *et al.* Exome sequencing of pleuropulmonary blastoma reveals frequent biallelic loss of TP53 and two hits in DICER1 resulting in retention of 5p-derived miRNA hairpin loop sequences. *Oncogene* **33**, 5295–5302 (2014).
45. Rakheja, D. *et al.* Somatic mutations in *DROSHA* and *DICER1* impair microRNA biogenesis through distinct mechanisms in Wilms tumours. *Nature Commun.* **2**, 4802 (2014).
46. Iwakawa, R. *et al.* Genome-wide identification of genes with amplification and/or fusion in small cell lung cancer. *Genes Chromosom. Cancer* **52**, 802–816 (2013).
47. Thériault, B. L., Dimaras, H., Gallie, B. L. & Corson, T. W. The genomic landscape of retinoblastoma: a review. *Clin. Experiment. Ophthalmol.* **42**, 33–52 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements G.Q.D. is supported by National Institutes of Health grant R01GM107536, Alex's Lemonade Stand Foundation, and the Ellison Medical Foundation. G.Q.D. is an affiliate member of the Broad Institute, and an investigator of the Howard Hughes Medical Institute and the Manton Center for Orphan Disease Research. J.T.P. was supported by Alex's Lemonade Stand Foundation. K.M.T. was supported as a Howard Hughes Medical Institute International Student Research Fellow and as a Herchel Smith Graduate Fellow. D.S.P. and R.E. were supported by award number T32GM007753 from the National Institute of General Medical Sciences.

Author Contributions G.Q.D. provided support and guidance for this work; G.Q.D. and J.T.P. conceived the hypothesis, designed the study, and wrote the manuscript; J.T.P. performed and interpreted most of the experiments and generated figures; K.M.T. helped perform the shRNA and siRNA experiments and generated most of the plasmid constructs; D.S.P. helped perform the RNA sequencing experiments; F.R., J.T., and F.B. generated the aCGH and survival data on neuroblastoma patients; C.S.S., K.C.K., and J.J.C. acquired tissue samples and assisted with the IHC analysis; R.E. assisted with the CRISPR experiments; M.S., Y.d.S., G.S.L.P., and S.J.R. provided technical help; P.C. processed RNA sequencing data and helped with data analysis; H.C.T. helped perform *in vitro* transfection experiments; A.H. assisted with RNA sequencing data analysis; J.K.O. and C.S.S. performed xenograft experiments; M.C. assisted with the RNA sequencing experiments.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.Q.D. (george.daley@childrens.harvard.edu).

METHODS

Cell culture. BE(2)C (ATCC CRL-2268), PA-1 (ATCC CRL-1572), IMR90 (ATCC CRL-186), SK-N-AS (ATCC CRL-2137), SH-SY5Y (ATCC CRL-2266), 293T (ATCC 11268), SK-N-DZ (ATCC CRL-2149), and Kelly cells (Sigma 92110411-1VL) were maintained in 1:1 DMEM/F12:MEM media with 10% inactivated fetal calf serum, $1\mu\text{g ml}^{-1}$ penicillin, and 1U ml^{-1} streptomycin. All cell lines were purchased for the purposes of this study, are not among commonly misidentified cell lines (according to the International Cell Line Authentication Committee), and tested negative for mycoplasma contamination.

Plasmids. *Turbo-RFP*, *LIN28B*, and *MYCN*(ORF) cDNAs were subcloned into the pcDNA3.1 expression vector (Invitrogen). The *MYCN* 3' UTR was cloned from BE(2)C cDNA and subcloned into pcDNA3.1:MYCN to create pcDNA3.1:MYCN/3' UTR. ShRNA-resistant-*LIN28B* and pcDNA3.1:MYCN/3' UTR-*let-7*-site-mutant vectors were made using a QuikChange site-directed-mutagenesis kit (Stratagene) on pcDNA3.1:*LIN28B* and pcDNA3.1:MYCN/3' UTR constructs, respectively. Wild-type and mutant *MYCN* 3' UTRs were subcloned into pcDNA3.1:RFP and psiCHECK2 to create pcDNA3.1:RFP/MYCN-3' UTRwt and pcDNA3.1:RFP/MYCN-3' UTRmut as well as psiCHECK2:MYCNwt and psiCHECK2:MYCNmut. pLS1:*DICER1*, pLS1:*IGF2BP1*, and pLS1:*Hmga2* were gifts from D. Bartel^{4,48} (Addgene plasmids 21649, 21639, and 14785). psiCHECK2-8x-*let-7* was a gift from Y. Tomari⁴⁹ (Addgene plasmid 20931).

siRNA/let-7 mimic transfections. BE(2)C and Kelly cells were reverse transfected using Lipofectamine 2000 (Life Technologies) into six-well plates using the appropriate siRNA or miRNA mimics (described below). Cells were harvested at time-points described for analysis by western blotting or qPCR. Growth assays were performed similarly, but in 96-well plates followed by time-point-specific BrdU growth assay. Global *let-7* target analysis was as follows: BE(2)C:MYCN-ORF cells co-transfected with control or MYCN-3' UTR-2 siRNA and either control or *let-7a* miRNA mimic were harvested 48 h after transfection. Control siRNA (Life Technologies 439846). *LIN28B* siRNAs were as follows: ORF1 (Life Technologies 4392420, identifier s52479), ORF2 (Life Technologies 4392420, identifier s52477), 5' UTR (GE Dharmacon Custom *LIN28B*-NM_001004317 Duplex siRNA, ON-TARGET Plus, sense: 5'-ACU GGA GAG AGG AGA GAA AUU-3', antisense: 5'-UUU CUC UCC UCU CUC CAG UUU-3'), 3' UTR-1 (GE Dharmacon J-028584-12-0020), 3' UTR-2 (GE Dharmacon Custom *LIN28B*-NM_001004317 Duplex siRNA, ON-TARGET Plus, sense: 5'-CAA CAG UGA UUG UGA GAA UUU-3', antisense: 5'-AUU CUC ACA ACU ACU GUU GUU-3'). MYCN siRNAs: ORF-1 (Life Technologies 4392420, identifier s9135), ORF-2 (Life Technologies 4392420, identifier s9134). Control miRNA mimic (Life Technologies 4464059), *let-7a* miRNA mimic (Life Technologies 4464066, identifier MC10050), *let-7a* inhibitor (Life Technologies 4464084, identifier MH10050).

Luciferase assays. For the MYCN 3' UTR loss of function assays, BE(2)C cells were reverse co-transfected using Lipofectamine 2000 in quadruplicate into 96-well plates with the appropriate luciferase vector and either control siRNA, MYCN siRNA, or *let-7a* mimic. Sixty hours after transfection, luciferase activity was assayed using the Dual Luciferase Reporter Assay System (Promega). For the MYCN 3' UTR gain of function assays, 293T cells were seeded into 96-well plates in quadruplicate and transfected using Lipofectamine 2000 the following day with the appropriate luciferase vector, MYCN 3' UTR overexpression vector, and either control miRNA mimic or *let-7a* mimic. Luciferase activity was measured 24 h after transfection as described above.

Immunohistochemistry. Immunohistochemistry was performed on human tumour tissue sections as previously described⁴ using anti-LIN28B antibody (Cell Signaling 4196) at a 1:400 dilution. Patient samples were obtained through Boston Children's Hospital IRB-CRS08-09-0429-2; Immunohistochemical and Molecular Analysis of Paediatric Tumours. Consent was obtained from all subjects.

Western blotting. Western blots were performed with antibodies against LIN28B (Cell Signaling 4196), LIN28A (Cell Signaling 3978), MYCN (Santa Cruz Biotechnology sc-53993), DICER1 Santa Cruz sc-30226), HK2 (Cell Signaling 2867S), IMP1 (Cell Signaling 2852S), tRFP (Origene TA150061), β -tubulin (Cell Signaling 2146), and β -actin (Santa Cruz Biotechnology sc-8342).

BrdU growth assay. Five thousand cells per well were plated on a 96-well plate in quadruplicate for each of three independent experiments. Cell proliferation was assayed relative to day 0 using a BrdU Cell Proliferation Assay Kit (Cell Signaling 6813) according to the manufacturer's protocol after incubation for 2 h with BrdU.

qPCR. Total RNA was isolated from cells using Trizol reagent (Life Technologies). For mRNA analysis, cDNA was prepared from $1\mu\text{g}$ RNA using Superscript II Reverse Transcriptase (Life Technologies) and random hexamers. Twenty nanograms of cDNA was then used for qPCR with the SYBR Select Master Mix (Life Technologies). mRNA primers: *LIN28B* (forward: GAG TCA ATA CGG GTA ACA GGA C; reverse: CAC CAC AGT TGT AGC ATC TAT CT) MYCN-1 (forward: CGA TTC AGA TGA TGA AGA TGA TGA AG; reverse: GAC AGC

CTT GGT GTT GGA), *IGF2BP1* (forward: CAG TCC AAG ATA GAC GTG CAT AG; reverse: CTC AGG GTT GTA AAG GGT AAG G), *DICER1* (forward: CTC CTA CCA CTA CAA TAC TAT CAC T; reverse: GGT CTT CAT AAA GGT GCT TGG T), β -actin (forward: GAC CCA GAT CAT GTT TGA GAC C; reverse: CGT AGC ACA GCT TCT CCT TAA T), *Hmga2* (forward: CTG CTC AGG AGG AAA CTG AAG; reverse: CAC TAA ACC TGG GAC TGT GAA G), *ARID3B* (forward: CAA GCA GAA TGG TGG TTT GG; reverse: ATG GAT GTG GGC AGG TTT AG). For *let-7* analysis, cDNA was prepared using 20 ng total RNA and a Taqman microRNA Reverse transcription kit (Life Technologies). Two nanograms of cDNA was then used for qPCR with the Taqman Universal PCR Master Mix (Life Technologies). Taqman microRNA Assays used (catalogue number 4427975): *let-7a* (identifier 000377), *let-7b* (identifier 002619), *let-7c* (identifier 000379), *let-7d* (identifier 002283), *let-7e* (identifier 002406), *let-7f* (identifier 000382), *let-7g* (identifier 002282), *let-7i* (identifier 002221), *miR-98* (identifier 000577), *U47* control microRNA Assay (identifier 001223). For both mRNA and *let-7* expression analysis, relative expression was determined using the $\Delta\Delta\text{CT}$ method, unless otherwise noted.

qPCR analysis of copies per cell. Total RNA per cell was determined by RNA harvest yield from 1.5×10^6 cells. For MYCN mRNA copies per cell, cDNA was prepared using 200 ng total RNA per cell type and from a dose curve of the synthetic MYCN-3' UTR-RNA-fragment (Integrated DNA Technologies) with Superscript III Reverse Transcriptase (Life Technologies) and a MYCN-gene-specific primer (MYCN-2 rev). Ten per cent of the cDNA reactions were then used for qPCR as described above. Linear regression analysis was performed on MYCN-3' UTR-RNA-fragment-sample CT values to determine the copies per reaction in the cell samples. MYCN mRNA copies per cell was then calculated on the basis of the total RNA per cell previously determined. Synthetic MYCN-3' UTR-RNA-fragment sequence was as follows: 5'-TTC CTA GCC TGT TTC CTC CTG TTA ATG TAT TTG TTC ATG TTT GGT GCA TAG AAC TGG GTA AAT GCA AAG TTC TGT GTT TAA TTT CTT CAA AAT GTA TAT ATT TAG TGC TGC ATC TTA TAG CAC TT-3'. MYCN-2 qPCR primers: forward: CCT AGC CTG TTT CTT CCT GTT A; reverse: GTG CTA TAA GAT GCA GCA CTA AAT. A similar strategy was used to determine *let-7* copies per cell. Synthetic *let-7* RNA molecules for each *let-7* family member (Integrated DNA Technologies) were used in dose curves for cDNA reactions for each *let-7*. cDNA was prepared as described above, and 10% of cDNA reactions were used for qPCR as described above. Linear regression analysis of the synthetic-*let-7* qPCR cycle threshold values and total RNA per cell values were used to calculate copies per cell for each *let-7* family member.

RNA-seq analysis. We analysed copies per cell as follows. Sequence reads were aligned to the human transcriptome using HISAT version 0.1.6 (ref. 50). The transcriptome file used for polyA-RNA-based libraries consisted of protein-coding and ncRNA sequences downloaded from ENSEMBL (release 80) ftp://ftp.ensembl.org/pub/release-80/fasta/homo_sapiens/, as well as the ERCC control sequences. The transcriptome file used for short-RNA-based libraries consisted of mature microRNA sequences downloaded from mirBase (release 21: <ftp://mirbase.org/pub/mirbase/>), in addition to the synthetic short RNA spike in oligonucleotides (small RNA spikes: LET7A2-5' SCR: /5Phos-rGrA rArGrA rUrGrU rGrGrU rGrUrU rGrUrA rUrArG rUrU; LET7I-5' SCR: 5'-Phos-rGrG rUrArG rArUrA rGrUrU rGrUrU rGrUrG rCrUrG rUrU; MIR-NEG: 5'-Phos-rUrU rArCrG rUrCrG rUrCrG rCrGrU rUrArU rU; MIR17-5' SCR: 5'-Phos-rArU rCrGrC rCrUrG rArUrA rArArG rUrGrC rArGrG rUrArG; JP_MIR-NEG DH1: 5'-Phos-rUrC rArCrA rArCrC rUrCrC rUrArG rArArA rGrArG rUrArG rA; JP_MIR-NEG DH2: 5'-Phos-rUrU rGrUrA rCrUrA rCrCrA rArCrA rArGrU rArGrU rGrG)⁵¹. To estimate transcript abundances, we applied Salmon version 0.3.2 (ref. 52) to the aligned reads and summarized transcript abundances into gene-level expression levels by summing all transcript expression levels mapping to the same gene. Gene-to-transcript mappings, and transcript type annotations, were downloaded also from ENSEMBL (ftp://ftp.ensembl.org/pub/release-80/gtf/homo_sapiens/). Unless otherwise stated, all RNA-seq data are presented as $\log_{10}(1 + \text{transcripts per million (TPM)})$, which we refer to as $\log(\text{TPM})$ henceforth. To estimate the absolute copy number of transcripts per cell, we performed linear regression of the spike-in oligonucleotides concentration (\log_{10}) on the $\log(\text{TPM})$ using the *lm* function in R for each sample. Using the slope and intercept estimated by this approach, we calculated the copies per cell of endogenous mRNA by determining the number of transcripts (on the basis of $\log(\text{TPM})$ values) present in the total nanograms of RNA used for library preparation. We then determined molecules per cell on the basis of pre-determined picograms of RNA per cell. Global *let-7* target analysis was as follows: polyA-selected mRNA sequencing reads were processed as above, with no spike in RNAs used. RNA-seq data sets are available under GEO reference series GSE81500 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE81500>).

Supplementary methodology: spike controlled quantification of small RNA. Through comparison with three miRNA spike-ins, including *let-7*-specific

5'-end-scrambled *let-7a* and *let-7i* spikes, we sought to minimize the effects of secondary structure bias known to exist during sRNA-seq library preparations, which can have significant effects on the relative efficiency of reads produced between different microRNAs⁵¹ (Extended Data Fig. 6). A total of six synthetic spikes were used. Disparate reads per million observed among the equimolar small RNA spikes (that is, the miR-Neg, LET7A2, LET7I spikes have much higher read counts than equimolar MIR17, miRNegDH1, and miRNegDH2 spikes) demonstrate both the need to use spikes similar to the miRNA-of-interest to accurately determine copies per cell and the risk inherent in using a single spike to determine copies per cell for all miRNAs. The use of multiple miRNA-specific spikes improves upon previous miRNA quantification strategies using a single control miRNA to quantify all miRNAs³¹ by limiting potential disparate read efficiencies between a single spike and the miRNA of interest. Difficulty in calculating relative expression of disparate miRNAs within a single data set because of such variable read efficiencies can therefore be mitigated through the use of spike-ins that closely resemble each miRNA of interest.

CRISPR/Cas9. Cas9/gRNA co-expressing lentiviral constructs (lentiCRISPRv2) were generated and lentiviral particles were produced as previously described^{53,54}, using protocols and gRNA design tools from <http://www.genome-engineering.org>. Puromycin selection began 24 h after lentiviral infection of BE(2)C and Kelly cells. Experiments were typically completed within 3 weeks of initial infection. Oligonucleotides used for gRNA cloning were as follows: *LIN28B* exon 2: CAC CGC ATC GAC TGG AAT ATC CAA G, AAA CCT TGG ATA TTC CAG TCG ATG C; *LIN28B* exon 3.1: CAC CGC AGA GCA AAC TAT TCA TGG A, AAA CTC CAT GAA TAG TTT GCT CTG C; *LIN28B* exon 3.2: CAC CGA ATG ATT ACC TAT CTC CCT T, AAA CAA GGG AGA TAG GTA ATC ATT C; *LIN28B* exon 4: CAC CGC CTT GTA GAT GCT ACA ACT G, AAA CCA GTT GTA GCA TCT ACA AGG C. Cas9/gRNA constructs: lentiCRISPRv2, lentiCRISPR:EGFPs-gRNA-1, and lentiCRISPR:EGFPs-gRNA-2 were gifts from F. Zhang^{53,54} (Addgene plasmids 52961, 51760, and 51761).

shRNA. Lentiviral particles were prepared as previously described¹¹. pLKO.1 short hairpin expression constructs (Sigma Mission shRNA): *LIN28B* shRNAs (sh1: TRCN0000144508, sequence 5'-CCTGTTTAGGAAGTGAAAGAA-3'; sh2: TRCN0000122599, sequence 5'-GCCTTGAGTCAATACGGGTAA-3'). Control vector (SHC001: no insert).

Xenografts. BE(2)C and SK-N-DZ cells were infected with either SHC001 or TRCN0000122599 pLKO.1 lentivirus, then puromycin selected for 48 h. One and a half million infected cells were injected subcutaneously into female Rag2 knockout (c57bl/6, 8-week-old) immune-deficient mice. Three weeks after injection, mice were killed and tumours were removed and weighed. This procedure was approved by the Boston Children's Hospital Institutional Animal Care and Use Committee under protocol 15-12-3071R, which limited xenograft tumour size to less than 2.0 cm.

aCGH analysis. Preparation of the 202 neuroblastoma aCGH data set has been previously described^{55,56}. MicroRNA-containing loci were analysed for gain or loss as previously described⁵⁷. Statistical significance between *MYCN*-amplified and

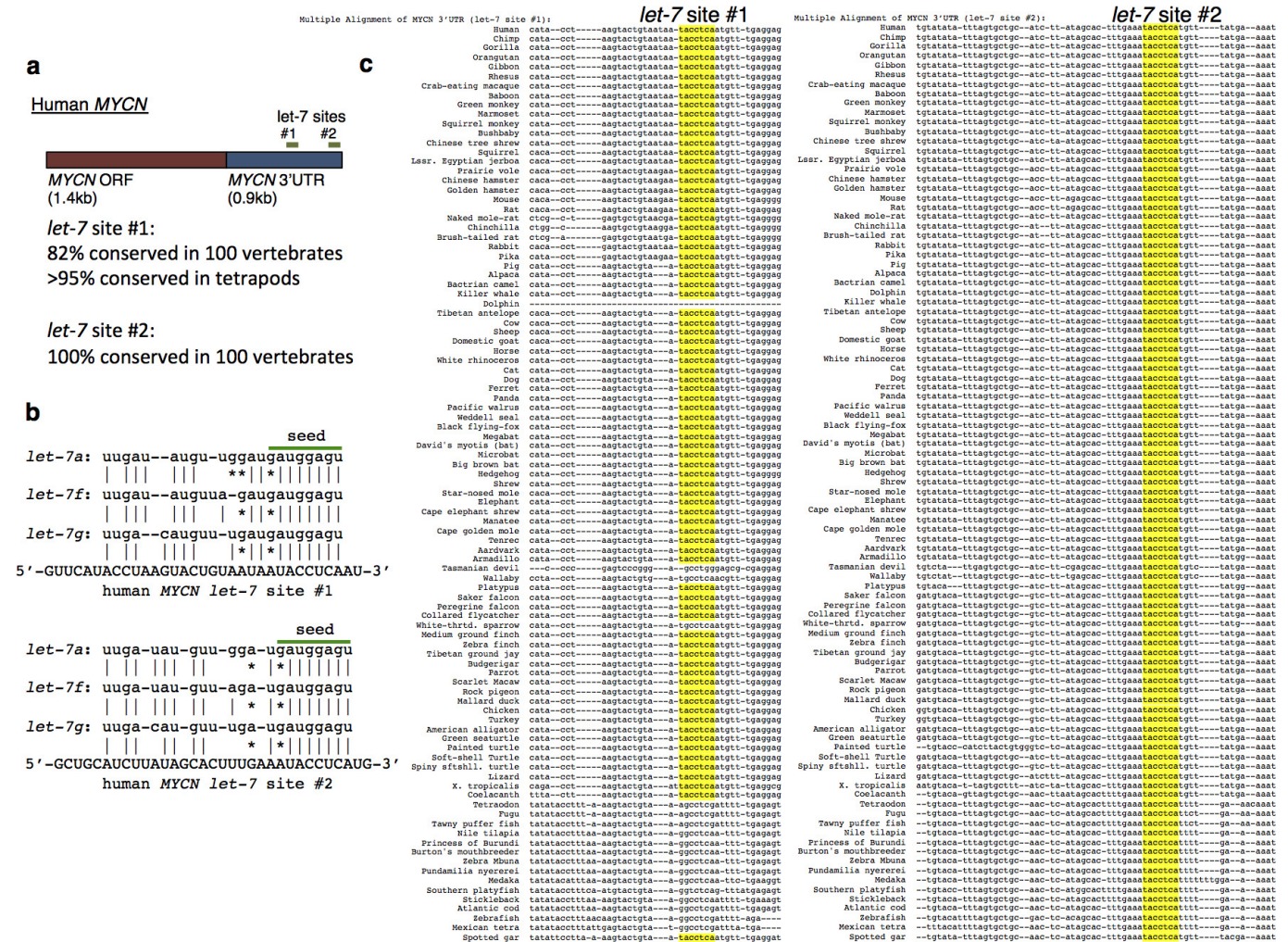
non-amplified tumours was determined using a *t*-test with Welch's adjustment on original copy number values. Kaplan-Meier curve generation and analysis was done using GraphPad Prism software.

ENCODE RNA-Seq data sets. Mature *let-7* expression data for 12 cell types were obtained from whole-cell small-RNA-Seq ENCODE/CSHL data sets, and *let-7* host transcript expression data from six cell types were obtained from RNA-seq ENCODE data sets on the University of California, Santa Cruz Genome Browser⁵⁸ (<http://genome.ucsc.edu/>). Expression levels were determined relative to *let-7c* and *let-7f* host transcript levels, respectively.

R2 database. Human neuroblastoma patient microarray and RNA-seq data sets were obtained from the R2: microarray analysis and visualization platform (<http://r2.amc.nl>) and analysed using GraphPad Prism software. Significance was determined by *t*-test with Welch's adjustment. Data sets used were Kozak (GEO accession number GSE45547) and SEQC (GEO accession number GSE62564).

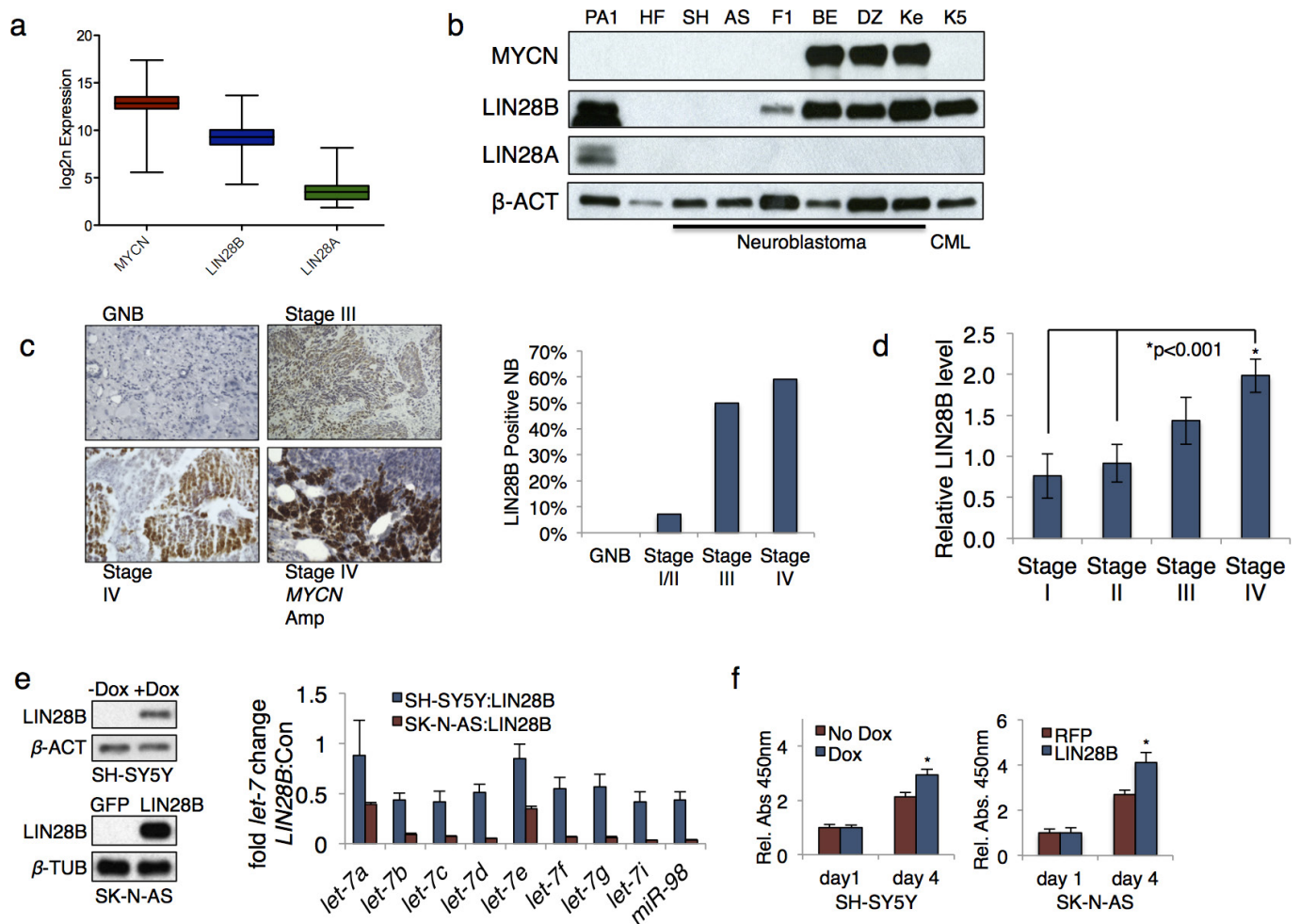
Statistical analysis. No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment. Unless otherwise noted, all experiments were performed at least three times independently. Statistical tests used are identified in each figure legend. *P* values less than 0.05 were considered significant.

48. Mayr, C. & Bartel, D. P. Widespread shortening of 3' UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**, 673–684 (2009).
49. Iwasaki, S., Kawamata, T. & Tomari, Y. *Drosophila* argonaute1 and argonaute2 employ distinct mechanisms for translational repression. *Mol. Cell* **34**, 58–67 (2009).
50. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* **12**, 357–360 (2015).
51. Hafner, M. *et al.* RNA-ligase-dependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries. *RNA* **17**, 1697–1712 (2011).
52. Patro, R., Duggal, G. & Kingsford, C. Salmon: accurate, versatile and ultrafast quantification from RNA-seq data using lightweight-alignment. *bioRxiv* <http://dx.doi.org/10.1101/021592> (2015).
53. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nature Methods* **11**, 783–784 (2014).
54. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014).
55. Spitz, R. *et al.* Oligonucleotide array-based comparative genomic hybridization (aCGH) of 90 neuroblastomas reveals aberration patterns closely associated with relapse pattern and outcome. *Genes Chromosom. Cancer* **45**, 1130–1142 (2006).
56. Thiessen, J. *et al.* Chromosome 17/17q gain and unaltered profiles in high resolution array-CGH are prognostically informative in neuroblastoma. *Genes Chromosom. Cancer* **53**, 639–649 (2014).
57. Rushlow, D. E. *et al.* Characterisation of retinoblastomas without RB1 mutations: genomic, gene expression, and clinical studies. *Lancet Oncol.* **14**, 327–334 (2013).
58. Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).

Extended Data Figure 1 | MYCN is a highly conserved *let-7* target.

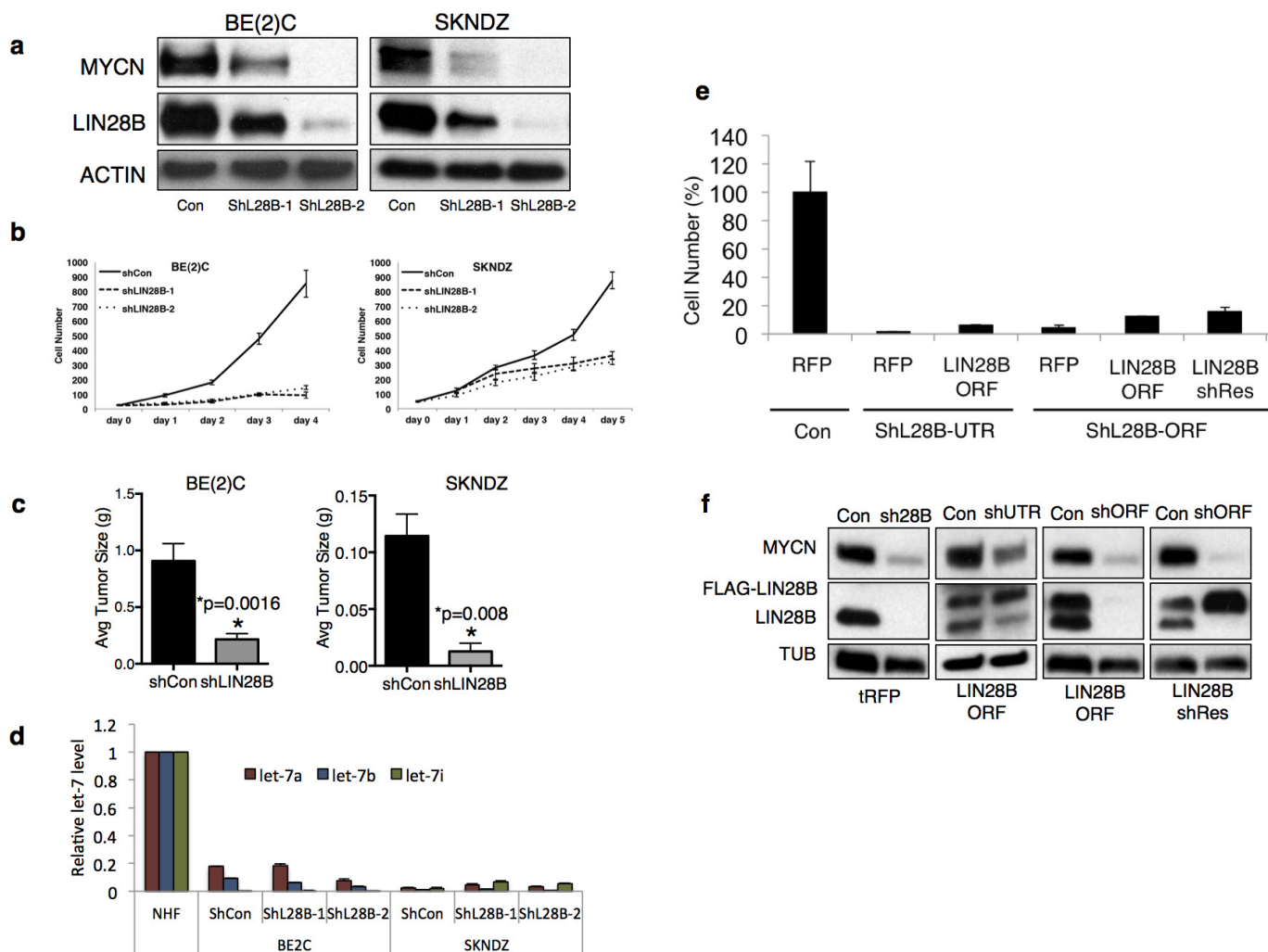
a, Schematic of human *MYCN* ORF and 3' UTR, indicating *let-7* sites 1 and 2 and their approximate location. **b**, Predicted base pairing patterns of *let-7a*, *let-7f*, and *let-7g* with *MYCN* *let-7* sites 1 and 2. A–G base pairs,

common in RNA, are represented by an asterisk. c, Alignments of *let-7* sites 1 and 2 in 100 vertebrate *MYCN* 3' UTRs (ENCODE, <https://genome.ucsc.edu/ENCODE/>).



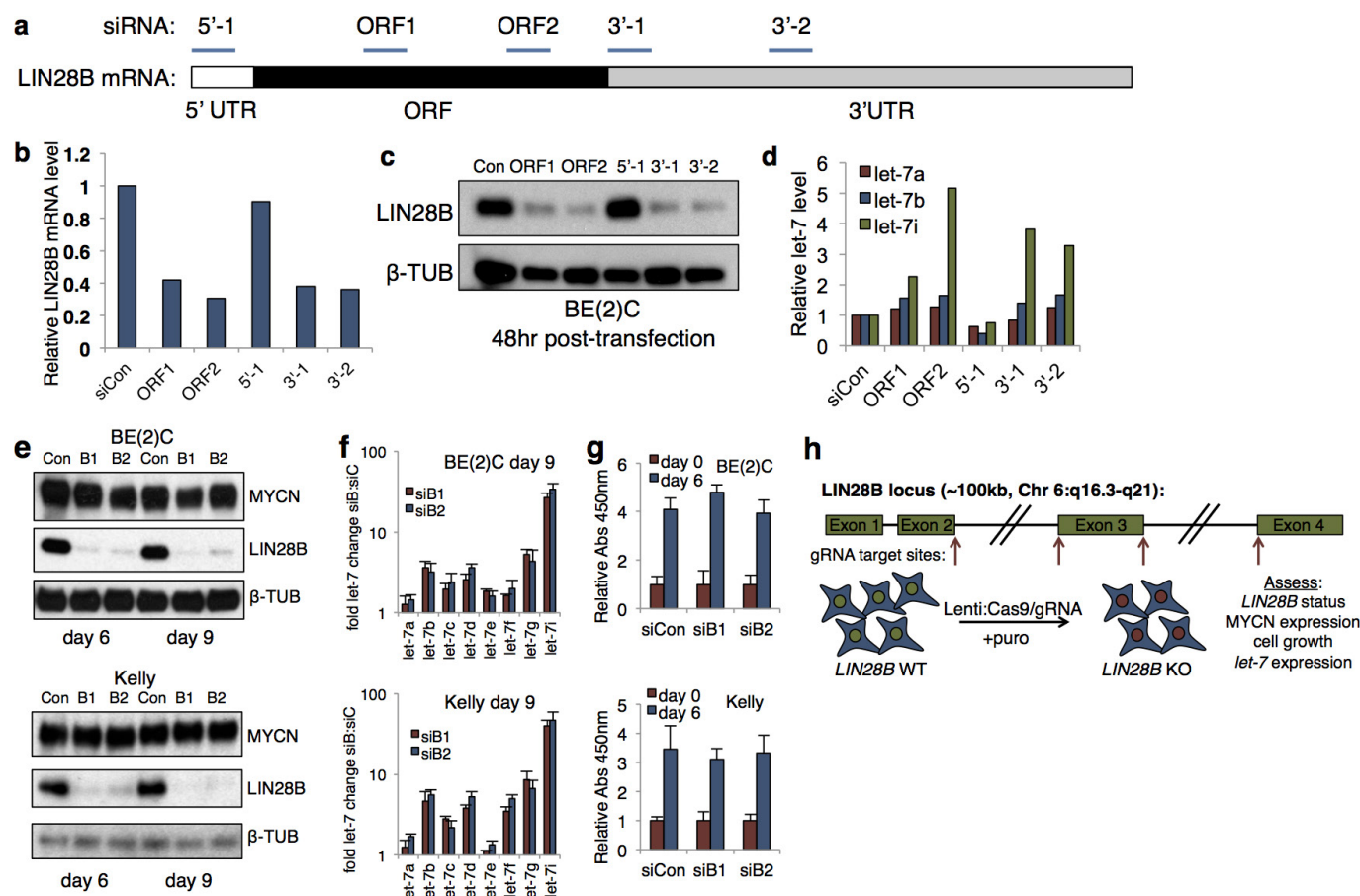
Extended Data Figure 2 | LIN28B expression and function in neuroblastoma. **a**, MYCN, LIN28B, and LIN28A mRNA expression levels in neuroblastoma ($n = 649$; see Source Data (ED Fig 2) in Supplementary Information). **b**, Immunoblot for indicated proteins in human embryonic carcinoma cells (PA1), normal human fibroblasts (HF), SK-N-SH (SH), SK-N-AS (AS), SK-N-F1 (F1), BE2C (BE), SK-N-DZ (DZ), Kelly (Ke), and human chronic myeloid leukaemia cells (K5). For gel source data, see Supplementary Figures. **c**, Representative LIN28B immunohistochemical staining of human neuroblastoma by stage

(left), percentage LIN28B positive neuroblastoma by disease stage (right); ($n = 36$). GNB, ganglioneuroblastoma. **d**, LIN28B expression by neuroblastoma stage ($n = 64$; Source Data (ED Fig 2)). **e**, Immunoblot for LIN28B in inducible LIN28B SH-SY5Y cells and GFP- or LIN28B-expressing SK-N-AS cells (left) and corresponding qPCR analysis of relative *let-7* family levels (right) (mean plus s.e.m. of three independent experiments shown). **f**, Relative growth rate (BrdU incorporation, right) of SH-SY5Y and SK-N-AS neuroblastoma cells from **d** (* $P < 0.05$, $n = 3$ independent experiments).



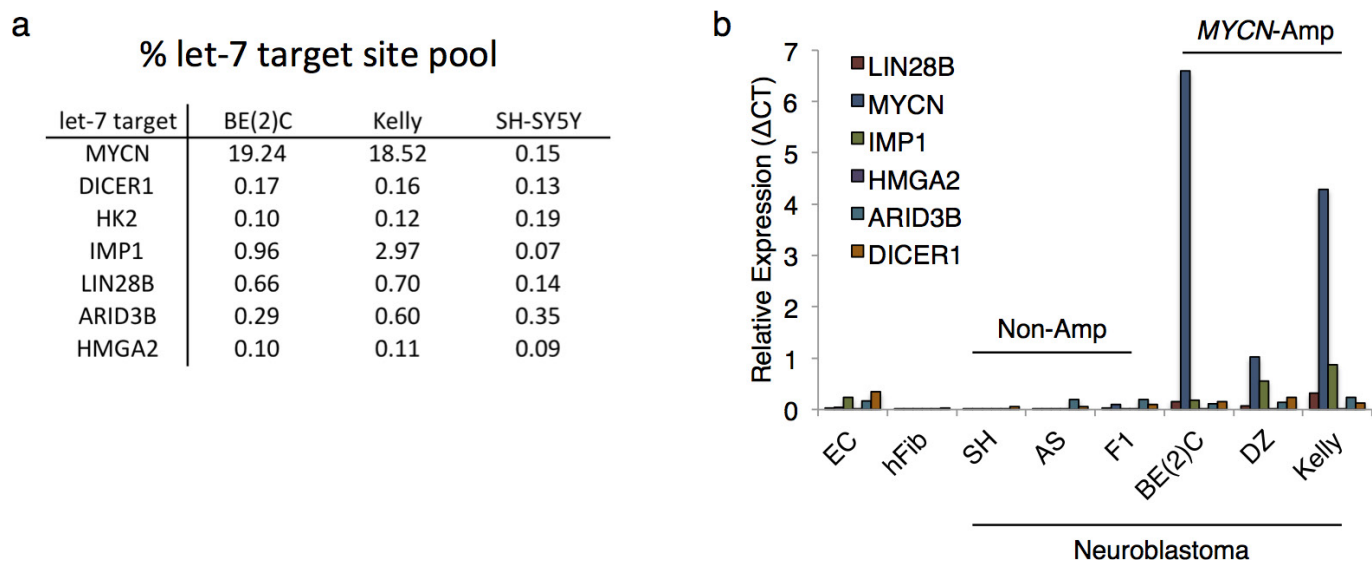
Extended Data Figure 3 | Short hairpin knockdown of *LIN28B* in neuroblastoma. **a**, Immunoblot for indicated proteins MYCN and LIN28B in MYCN-amplified cells infected with *LIN28B* targeting lentiviral shRNAs. For gel source data, see Supplementary Figures. **b**, Cell proliferation analysis of cells described in **a** ($n=3$ independent experiments). **c**, Average tumour size of human-mouse subcutaneous xenograft tumour analysis 3 weeks after injection of 2×10^6 cells infected with a *LIN28B* targeting lentiviral shRNA ($n=6$ mice for BE(2)C, $n=3$ mice for SK-N-DZ; Supplementary Figures and Source Data (ED Fig 3)). **d**, qPCR analysis of *let-7a*, *let-7b*, and *let-7i* levels in cells described

in **a** (mean plus s.e.m. of three independent experiments shown). **e**, Cell proliferation analysis of BE(2)C cells stably expressing red fluorescence protein (RFP), Flag-tagged *LIN28B* ORF, or shRNA resistant Flag-tagged *LIN28B* (LIN28B shRes) infected with *LIN28B* lentiviral shRNAs targeting the *LIN28B* 3' UTR (ShL28B-UTR) or the *LIN28B* open-reading frame (ShL28B-ORF). Cell counts were performed 7 days after lentiviral shRNA infection (mean plus s.e.m. of three independent experiments shown). **f**, Immunoblot for indicated proteins in cells described in **e**. For gel source data, see Supplementary Figures.

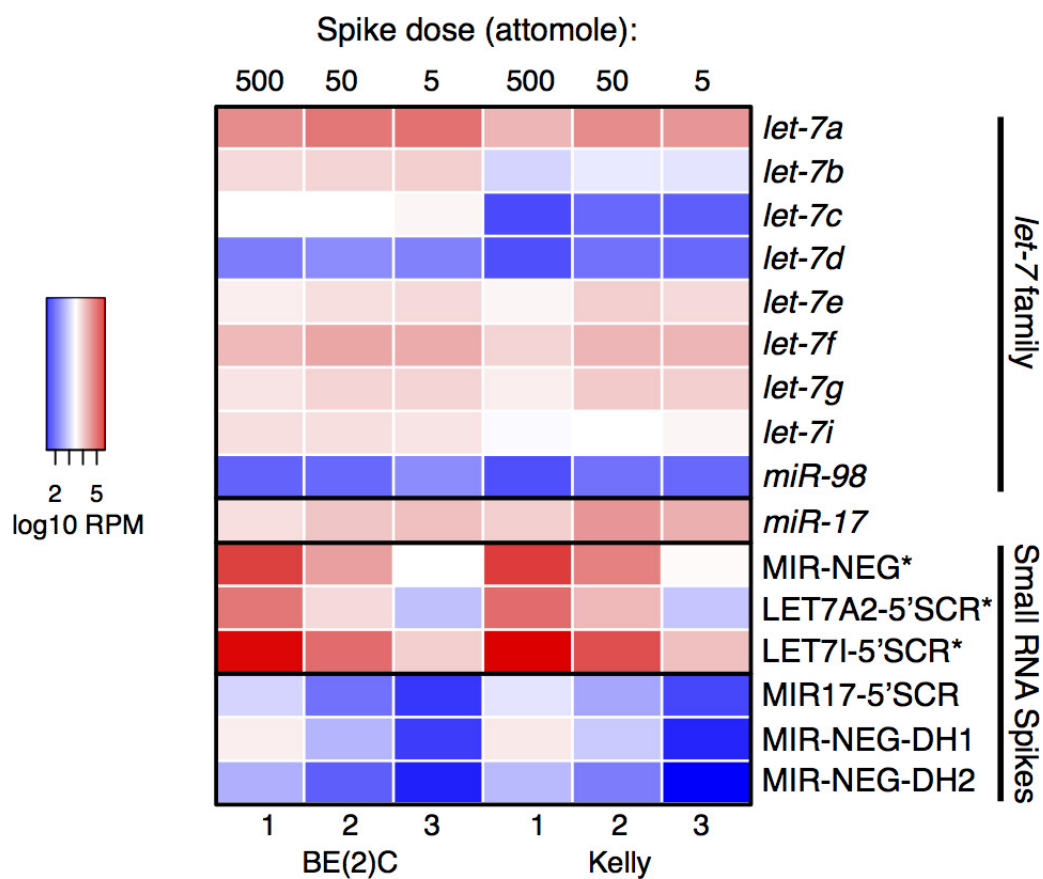


Extended Data Figure 4 | Small interfering RNA knockdown of *LIN28B* in neuroblastoma. **a**, Schematic of approximate siRNA target sites within the *LIN28B* mRNA. **b**, qPCR analysis of *LIN28B* mRNA levels in BE(2)C cells 48 h after transfection with the indicated *LIN28B* targeting siRNAs (mean of two independent experiments shown). **c**, Immunoblot analysis of *LIN28B* in cells from **a**. For gel source data, see Supplementary Figures. **d**, qPCR analysis of indicated *let-7* levels in cells from **a** (mean of two independent experiments shown). **e**, Immunoblot analysis of MYCN

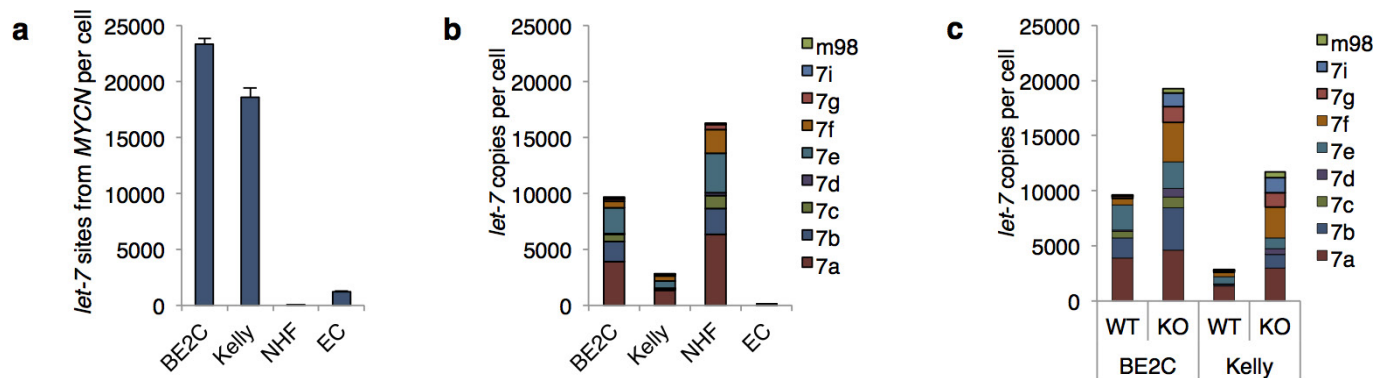
and *LIN28B* in serially transfected MYCN-amplified cells for 6 or 9 days. Identical transfections were performed on days 0, 3, and 6. For gel source data, see Supplementary Figures. **f**, Day 9 qPCR analysis of the *let-7* family in the cells from **a** ($n = 3$ independent experiments, mean plus s.e.m. shown). **g**, Cell growth analysis of day 0 to day 6 cells from **a** (BrdU incorporation, $n = 3$ independent experiments, mean plus s.e.m. shown). **h**, Lentiviral CRISPR-Cas9/*LIN28B* gRNA strategy targeting *LIN28B* at four distinct exon/intron junctions used in **b–g**.



Extended Data Figure 5 | Relative levels of *let-7* targets in neuroblastoma. **a**, mRNA-seq *let-7* target table (as percentage *let-7* target-site pool). **b**, qPCR analysis of indicated *let-7* targets in neuroblastoma cells, PA1 embryonic carcinoma cells (EC), and normal human fibroblasts (hFib). Expression relative to β -ACTIN (Δ CT method) (mean of two biological replicates shown).

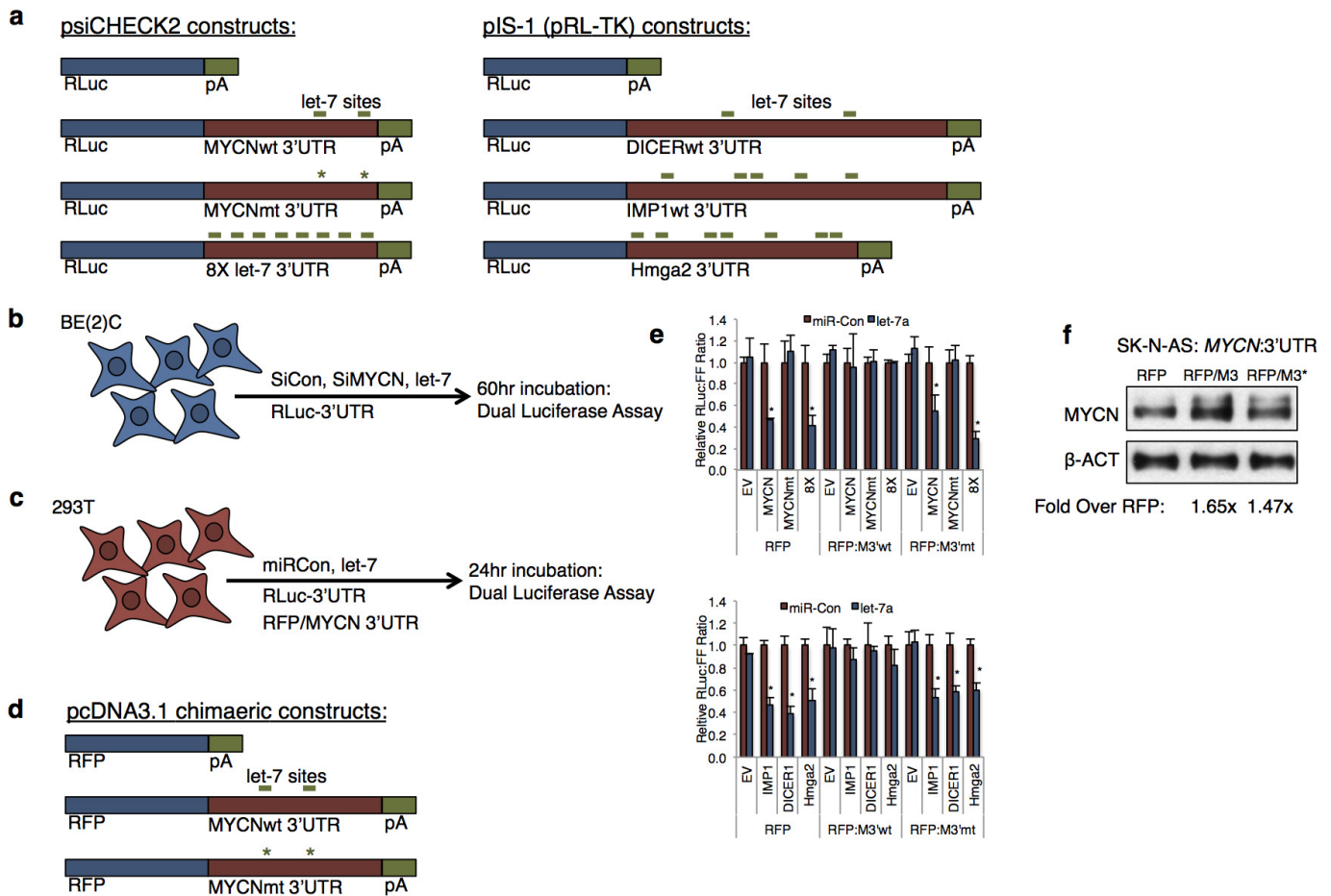


Extended Data Figure 6 | Heat map of *let-7* and small RNA spike reads. Heat map of three BE(2)C and three Kelly sRNA-seq samples depicting the relative reads per million of the *let-7* family, *miR-17*, and the six small RNA spikes added in equimolar amounts per sample (spikes miR-Neg, LET7A2, and LET7I were used to determine *let-7* copies per cell from the small RNA sequencing data set). RPM, reads per million.



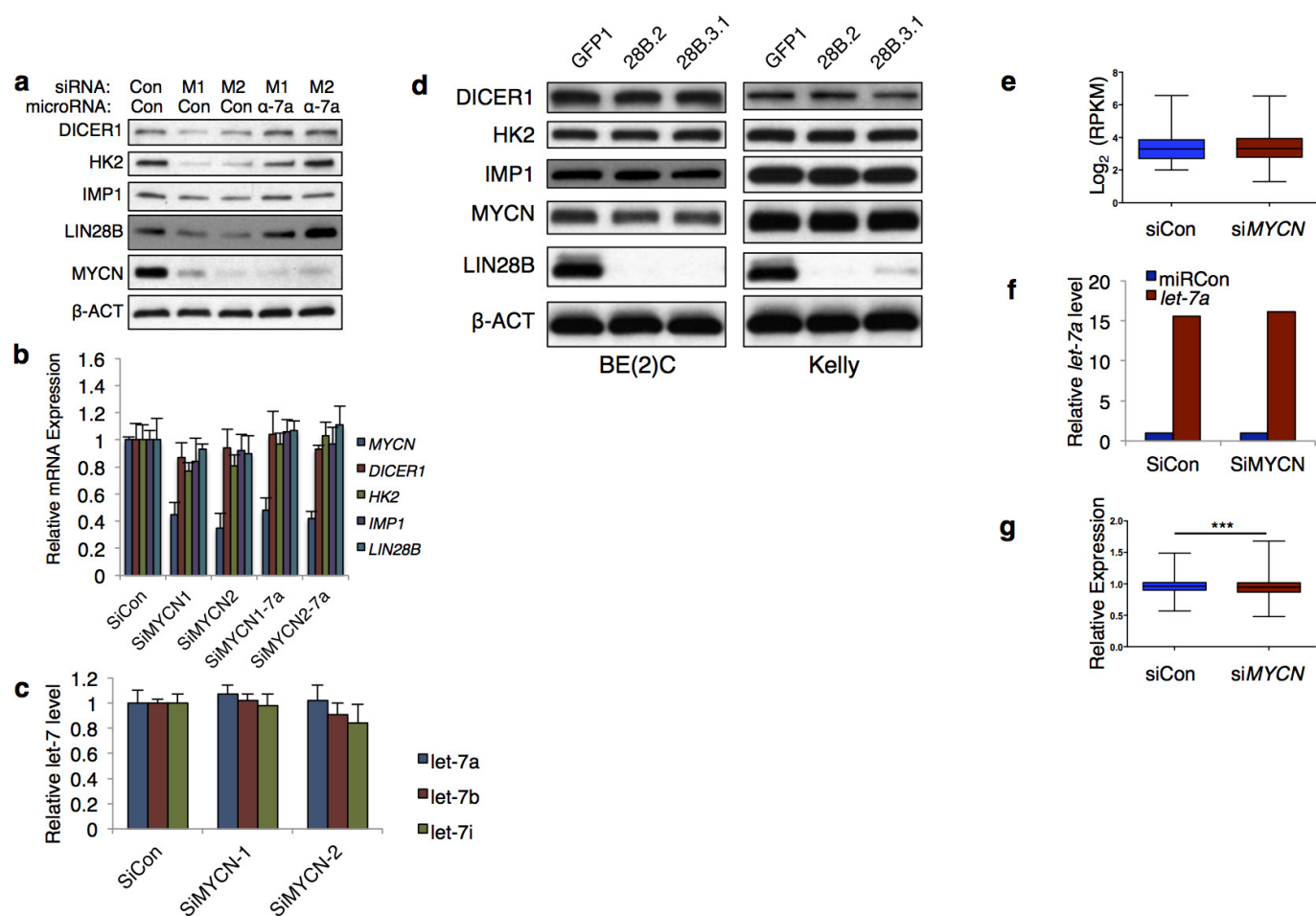
Extended Data Figure 7 | qPCR quantification of *MYCN* and *let-7* copies per cell. **a**, Total *let-7* sites per cell provided by *MYCN* mRNA in BE(2)C, Kelly, normal human fibroblasts (NHF), and embryonic carcinoma cells (EC) (mean plus s.e.m. of three biological replicates shown). **b**, Total *let-7* copies per cell in cells from **a**, presented as stacked

graphs of all *let-7* family members (mean of three biological replicates shown). **c**, Total *let-7* copies per cell in wild-type or *LIN28B* knockout BE(2)C and Kelly cells, presented as stacked graphs of all *let-7* family members (values derived from *let-7* copies per cell in **b** and average *let-7* fold change described in Fig. 2f, g).



Extended Data Figure 8 | Luciferase reporter and gain of function constructs. **a**, Luciferase constructs used in the luciferase assays in Fig. 3d and Extended Data Fig. 8e. **b**, Schematic of the luciferase transfection protocols used in Fig. 3d. **c**, Schematic of the luciferase protocol used in Extended Data Fig. 8e. **d**, pcDNA3.1 constructs used in Extended Data Fig. 8e, f. **e**, Top: relative luciferase ratio in 293T cells co-transfected with the indicated 3' UTR luciferase and pcDNA3.1 vectors in the presence of either control miRNA or *let-7a* mimic.

Bottom: relative luciferase ratio in 293T cells co-transfected with the indicated 3' UTR luciferase and pcDNA3.1 vectors in the presence of either a control miRNA or *let-7a* mimic. Mean of four independent experiments plus s.e.m. shown (* $P < 0.05$ relative to empty vector, unpaired *t*-test). **f**, Immunoblot analysis of MYCN in SK-N-AS cells stably expressing a MYCN ORF + 3' UTR transgene and transfected with the indicated pcDNA3.1 vector. For gel source data, see Supplementary Figures.



Extended Data Figure 9 | MYCN mRNA sponges *let-7*. **a**, Immunoblot analysis of indicated proteins in BE(2)C cells transfected for 2.5 days with control, *MYCN-1* (M1), or *MYCN-2* (M2) siRNA and either control microRNA or *let-7a* inhibitor. For gel source data, see Supplementary Figures. **b**, qPCR analysis of *DICER1*, *HK2*, *IMP1*, *LIN28B*, and *MYCN* in cells transfected as in **a**. **c**, qPCR analysis of *let-7a*, *let-7b*, and *let-7i* in BE(2)C cells transfected for 2.5 days with control siRNA, siM1, or siM2 ($n = 3$ independent experiments, mean plus s.e.m. shown). **d**, Immunoblot analysis of indicated proteins in cells infected with indicated Cas9-gRNA lentivirus. For gel source data, see Supplementary Figures. **e**, Expression

levels of *let-7* targets in BE(2)C:MYCN cells transfected with siCon or siMYCN-3' UTR. **f**, Relative *let-7* expression in BE(2)C:MYCN cells co-transfected with siCon or siMYCN (3' UTR) siRNA and miRCon or *let-7a* mimic. A 16-fold increase in total *let-7*, owing to *let-7a* making up almost half of the total cellular pool (Fig. 3c, lower). **g**, Relative expression levels of *let-7* targets in siCon and siMYCN cells transfected with *let-7a* mimic (data represent one round of mRNA-seq, *** $P < 0.001$, one-tailed Wilcoxon test, GSE81497, see Source Data F3).

a % incidence in neuroblastoma patients

	3p21 loss	11q24 loss	3p/11q loss	MYCN Amp
Pugh, Morozova (n=240)	32.5	47.9	27.5	32.9
Spitz (n=90)	38.9	57.8		21.1
Maris (n=295)		42.0		31.8
Nair (n=86)	24.4	43.0	23.3	30.2

b Mature *let-7* Level Analysis (sRNA Seq)

Name	Type	GEO Accession
IMR90	Lung Fibroblasts	GSM973680
HAoAF	Aortic Adventitial Fibroblasts	GSM977035
HAoEC	Aortic Endothelial Cells	GSM977040
B-CD20+	B Cells	GSM977043
Mono-CD14+	Monocytes	GSM977044
HMSC-BM	Undifferentiated BM MSC	GSM977028
NHEK	Epidermal Keratinocytes	GSM897086
NHEM M2	Epidermal Melanocytes	GSM977047
HPC	Undifferentiated Pericytes	GSM977049
SKNSH	Non-Amp Neuroblastoma	GSM973667
HeLa	Cervical Cancer	GSM897079
MCF7	Breast Adenocarcinoma	GSM897081

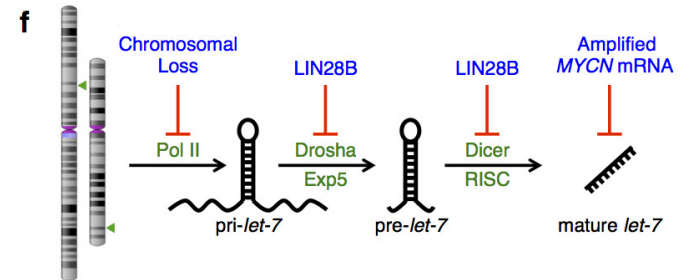
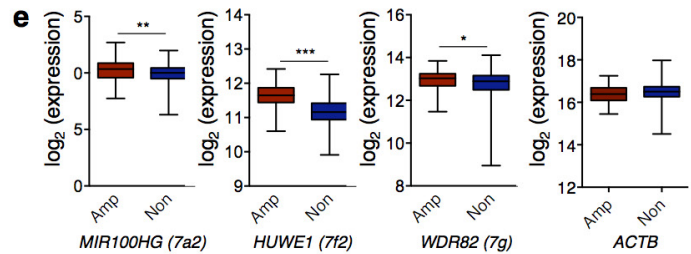
<i>let-7</i> locus	Host Transcript	Transcript Class	<i>let-7</i> location
7a1, 7d, 7f1	LET7DHG	lncRNA	intron
7a2, miR100	MIR100HG	lncRNA	intron
7a3, 7b	LET7BHG	lncRNA	exon
7c, miR99a	MIR99AHG	lncRNA	intron
7e, miR99b	SPACA6P-AS	lncRNA	exon
7f2, miR98	HUWE1	protein coding	intron
7g	WDR82	protein coding	intron
7i	ENST00000550290	predicted lncRNA	intron

Extended Data Figure 10 | Neuroblastoma patient and ENCODE data.

a, Detail of the incidence of chromosome 3p21 and 11q23 loss and *MYCN* amplification as determined by analysis of the indicated retrospective chromosomal aberration studies on neuroblastoma. **b**, List of the ENCODE sRNA-seq samples analysed (with associated GEO accession numbers) for the relative expression of mature *let-7* in Fig. 5c. **c**, List of *let-7* family host transcripts, transcript class, and *let-7* location within the transcript. **d**, List of the ENCODE mRNA-seq samples analysed (with associated University of California, Santa Cruz submission identifier

d Host-Transcript Expression Analysis (mRNA Seq)

Name	Type	UCSC Submission ID
NHLF	Lung Fibroblasts	3413
NHEK	Epidermal Keratinocytes	3473
HUVEC	Umbilical Vein Endothelial Cells	3470
HSMM	Skeletal Muscle Myoblasts	3468
H1-ESC	Embryonic Stem Cells	3464
HeLa	Cervical Carcinoma	3418



numbers) for the relative expression of *let-7* host transcripts in Fig. 5d.

e, Relative expression of *let-7a2*, *let-7f2*, and *let-7g* host genes by microarray in *MYCN*-amplified and non-amplified neuroblastoma. *ACTB* shown as control. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, unpaired t -test, $n = 643$, Source Data (ED Fig 10). **f**, Schematic showing the several mechanisms that impair *let-7* biogenesis and function in neuroblastoma (chromosome images created at <http://www.ncbi.nlm.nih.gov/genome/tools/gdp/>).

A novel cereblon modulator recruits GSPT1 to the CRL4^{CRBN} ubiquitin ligase

Mary E. Matyskiela^{1*}, Gang Lu^{1*}, Takumi Ito^{2*}, Barbra Pagarigan¹, Chin-Chun Lu¹, Karen Miller¹, Wei Fang¹, Nai-Yu Wang¹, Derek Nguyen¹, Jack Houston¹, Gilles Carmel¹, Tam Tran¹, Mariko Riley¹, Lyn'Al Nosaka³, Gabriel C. Lander³, Svetlana Gaidarova¹, Shuichan Xu¹, Alexander L. Ruchelman¹, Hiroshi Handa², James Carmichael¹, Thomas O. Daniel¹, Brian E. Cathers¹, Antonia Lopez-Girona¹ & Philip P. Chamberlain¹

Immunomodulatory drugs bind to cereblon (CRBN) to confer differentiated substrate specificity on the CRL4^{CRBN} E3 ubiquitin ligase. Here we report the identification of a new cereblon modulator, CC-885, with potent anti-tumour activity. The anti-tumour activity of CC-885 is mediated through the cereblon-dependent ubiquitination and degradation of the translation termination factor GSPT1. Patient-derived acute myeloid leukaemia tumour cells exhibit high sensitivity to CC-885, indicating the clinical potential of this mechanism. Crystallographic studies of the CRBN–DDB1–CC-885–GSPT1 complex reveal that GSPT1 binds to cereblon through a surface turn containing a glycine residue at a key position, interacting with both CC-885 and a ‘hotspot’ on the cereblon surface. Although GSPT1 possesses no obvious structural, sequence or functional homology to previously known cereblon substrates, mutational analysis and modelling indicate that the cereblon substrate Ikaros uses a similar structural feature to bind cereblon, suggesting a common motif for substrate recruitment. These findings define a structural degron underlying cereblon ‘neosubstrate’ selectivity, and identify an anti-tumour target rendered druggable by cereblon modulation.

The immunomodulatory drugs lenalidomide and pomalidomide promote recruitment and ubiquitination of substrate proteins to the CRL4^{CRBN} (CUL4–DDB1–RBX1–CRBN) E3 ubiquitin ligase, leading to their subsequent degradation by the 26S proteasome^{1,15,16}. The therapeutic benefits of immunomodulatory drug treatment in multiple myeloma and 5q-deletion-associated myelodysplastic syndrome are driven by a combination of both anti-proliferative cellular and immunomodulatory effects^{2–5}. The CRL4^{CRBN} E3 ubiquitin ligase complex consists of a core scaffolding component CUL4, which binds to the RBX1 subunit, the docking site for E2 recruitment, and DDB1, an adaptor protein. DDB1 provides a binding site for CRL4 substrate receptor proteins such as cereblon, termed DCAFs (DDB1 and CUL4 associated factors), which recruit substrates to the ubiquitin ligase complex and determine substrate specificity^{6–9}.

Structural studies have shown that immunomodulatory drugs bind in a shallow hydrophobic pocket on the surface of cereblon^{10,11}. Cereblon binding is mediated by a glutarimide ring, a feature common to the clinical molecules thalidomide, lenalidomide, pomalidomide, and CC-122. These compounds feature additional chemical groups that are not contained by the cereblon binding pocket, and have been postulated to interact with cognate substrates resulting in ‘neomorphic’ E3 ligase activity. Several proteins have been reported to be ubiquitinated by the CRL4^{CRBN} E3 ligase upon treatment with immunomodulatory drugs. Ikaros and Aiolos are zinc finger transcription factors important in haematological differentiation^{12,13}, which are degraded by addition of several immunomodulatory drugs^{5,14–16}. Recently, lenalidomide, but not pomalidomide or CC-122, was demonstrated to induce the degradation of the protein kinase casein kinase 1 α (CK1 α), thereby exploiting CK1 α haploinsufficiency associated with 5q-deletion-associated myelodysplastic syndrome⁵. Importantly, these cereblon substrates share no obvious sequence or structural homology to explain their susceptibility to immunomodulatory drugs.

We investigated whether cereblon could be repurposed to induce the degradation of previously unidentified targets resulting in anti-tumour activity. We identified CC-885 as a potent anti-cancer agent eliciting broad spectrum growth inhibition against cancer cell lines and patient-derived acute myeloid leukaemia (AML) cells, indicating the clinical potential for this mechanism. We identified GSPT1 (eRF3a) as a novel CC-885-dependent cereblon substrate mediating the anti-proliferative effects of CC-885. GSPT1 is a translation termination factor that binds eRF1 to mediate stop codon recognition and nascent protein release from the ribosome^{17–19}. GSPT1 is not targeted for degradation by any of the current generation of clinically approved immunomodulatory drugs, and possesses no obvious homology to Ikaros, Aiolos or CK1 α . To investigate the structural determinants for substrate recruitment, we solved the structure of cereblon in complex with DDB1, CC-885, and GSPT1. We found that the main contacts to cereblon–CC-885 by GSPT1 are mediated by the peptide backbone of a surface turn. Degrons, molecular motifs that direct E3 ligase interactions, have been defined for other ligases^{20–22}, and include examples where recognition is dependent upon small molecules binding to the ligase complex^{23,24}. Mutational evidence indicates that Ikaros is recruited via a motif similar to that of GSPT1, demonstrating that there is a common structural degron for ligand-directed recruitment to the CRL4^{CRBN} E3 ligase.

CC-885 is a cereblon-dependent anti-tumour agent

On the basis of the hypothesis that additional proteins are vulnerable to cereblon-mediated degradation, we sought to identify cereblon ligands with unique activities from a library of analogues. CC-885 was identified as a compound exhibiting a strong anti-proliferative phenotype in tumour cell lines (Extended Data Fig. 1a). The broad activity profile of CC-885 is highly differentiated from that of thalidomide, lenalidomide and pomalidomide (Extended Data Fig. 1a, d). To assess the potential

¹Celgene Corporation, 10300 Campus Point Drive, Suite 100, San Diego, California 92121, USA. ²Department of Nanoparticle Translational Research, Tokyo Medical University, Shinjuku-ku, Tokyo 160-8402, Japan. ³The Scripps Research Institute, San Diego, California 92121, USA.

*These authors contributed equally to this work.

clinical utility of CC-885, the molecule was tested in an *ex vivo* setting to evaluate potency against patient-derived AML tumour cells. CC-885 exhibited sub-nanomolar potency against 4 out of 5 patient samples (Extended Data Fig. 1b), with reduced activity against the normal lymphoid cells obtained from the same donors.

We next established the cellular target of CC-885. Like lenalidomide and its analogues, CC-885 contains a glutarimide ring that can directly interact with cereblon (Fig. 1a). The anti-proliferative effects of CC-885 are cereblon-dependent, as ablation of cereblon gene expression using clustered regularly interspaced short palindromic repeats (CRISPR) conferred resistance to CC-885 in 293FT human embryonic kidney cells, as well as in the AML cell lines NB-4, MOLM-13 and OCI-AML2 (Fig. 1b, Extended Data Fig. 1c). To identify potential substrates of CRL4^{CRBN} regulated by CC-885, we generated 293T HEK cells stably expressing Flag- and haemagglutinin (HA)-double-tagged cereblon (FH-CRBN). Using large-scale anti-Flag affinity purification followed by mass spectrometry analysis, GSPT1 (eRF3a) was identified as a protein that associated with FH-CRBN specifically in the presence of CC-885 (Fig. 1c, Supplementary Information Table 1). The presence of GSPT1 in the anti-Flag immunoprecipitates of FH-CRBN complex bound to CC-885 was confirmed by immunoblot analysis (Fig. 1c, bottom). Reciprocal binding assays performed with HA-tagged GSPT1 immobilized on anti-HA resin further supported that CC-885 could enhance the binding affinity of endogenous cereblon for GSPT1 (Extended data Fig. 2). Lenalidomide did not promote the interactions between cereblon and GSPT1. Conversely, CC-885 enhanced the binding of Ikaros (IKZF1) to cereblon. Consistent with the binding data, CC-885, but not lenalidomide or pomalidomide, triggered the depletion of GSPT1, whereas all three agents decreased the level of Ikaros in NB-4 leukaemia cells (Fig. 1d).

CC-885 promotes cereblon-dependent GSPT1 destruction

We next determined whether the changes in GSPT1 protein levels upon CC-885 treatment resulted from CRL4^{CRBN}-dependent ubiquitination and degradation. In 293FT HEK cells, the CC-885-induced reduction of GSPT1 protein levels can be prevented by depletion of cereblon via CRISPR or by co-treatment with either a neddylation inhibitor (MLN4924) or a proteasome inhibitor (MG132) (Extended Data Fig. 3, and data not shown). The mRNA level of GSPT1 was not decreased but rather increased in 293FT HEK cells treated with CC-885, consistent with a non-transcriptional mechanism (Extended Data Fig. 3). Together, these results indicate that CC-885 treatment causes cereblon-dependent ubiquitination and subsequent proteosomal degradation of GSPT1.

CC-885 promoted both the *in vivo* and *in vitro* ubiquitination of GSPT1 by cereblon (Fig. 1e, f and Extended Data Fig. 4) and decreased the protein half-life of GSPT1 in 293FT HEK parental but not in *CRBN*^{-/-} cells (Extended Data Fig. 4). Reintroduction of wild-type cereblon via lentiviral infection in 293FT *CRBN*^{-/-} cells restored the CC-885-dependent degradation of GSPT1, whereas CRBN(W386A) and CRBN(E377V) mutants displayed diminished activity, with E377V exhibiting a stronger effect consistent with the critical role of E377 in anchoring CC-885 in the cereblon-GSPT1 complex (Extended Data Fig. 4).

GSPT1 depletion drives CC-885 effects

To address whether the cereblon-dependent degradation of GSPT1 was responsible for the cytotoxic effects of CC-885, a GSPT1 mutant that retains its normal function, but loses CC-885-dependent cereblon binding, was used to distinguish the role of GSPT1 from that of other substrates. The yeast homologue, SUP35, shares a high degree of

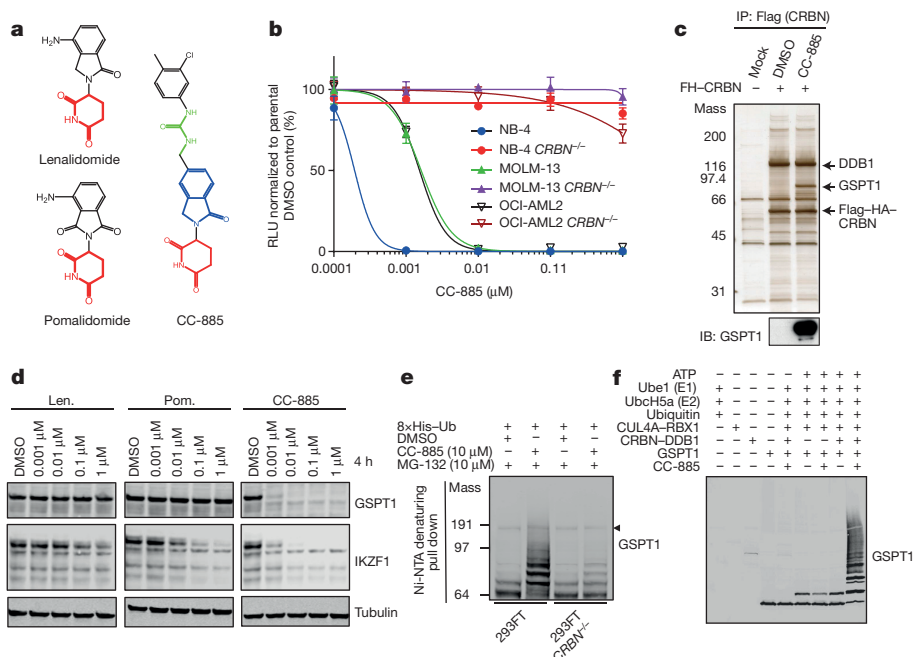


Figure 1 | GSPT1 is a substrate of the cereblon E3 ligase complex bound with CC-885. **a**, Structures of lenalidomide, pomalidomide and CC-885 with the glutarimide ring shown in red, isoindolinone in blue, the methylene urea in green. **b**, The effect of CC-885 on cell proliferation in parental and *CRBN*^{-/-} AML cell lines as indicated. Result is representative of three biological replicates. Data are shown as mean \pm s.d., $n = 3$ technical replicates. **c**, Silver stained gel of the immunoprecipitation of Flag-tagged cereblon to identify GSPT1 binding in the presence of CC-885. **d**, Immunoblot analysis of whole-cell extracts of NB-4 cells incubated with DMSO, or with lenalidomide, pomalidomide or CC-885 at the concentrations indicated for 4 h. **e**, The *in vivo* ubiquitination of GSPT1 is cereblon dependent. 293FT *CRBN*^{-/-} cells were transfected with

plasmids expressing 8 \times His-Ub. After 48 h, cells were treated with CC-885 and MG-132 as indicated for 8 h. Whole-cell extract and ubiquitinated protein products enriched with nickel sepharose were subjected to immunoblot analysis. The anti-His ubiquitin immunoblot showing equal enrichment of ubiquitinated proteins is shown in Extended Data Fig. 3a. Arrowhead indicates nonspecific band. **f**, *In vitro* ubiquitination of GSPT1 by the CUL4A-RBX1-DDB1-cereblon complex. Recombinant protein products as indicated were incubated with or without ATP (10 mM) and CC-885 (100 μM) in the ubiquitination assay buffer at 30°C for 2 h, and then analysed by immunoblotting. Results in **d-f** are representative of three independent experiments. For gel source data, see Supplementary Information Fig. 1.

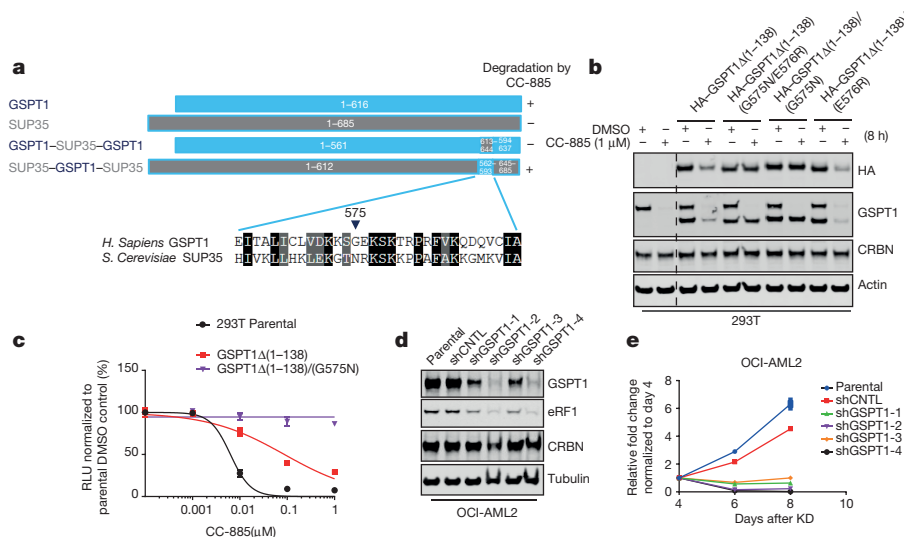


Figure 2 | GSPT1 degradation is necessary and sufficient for the CC-885-associated cytotoxicity. **a**, Schematic of the human GSPT1 and *Saccharomyces cerevisiae* SUP35 chimaeras. The CC-885-dependent degradation motif of human GSPT1, as well as the corresponding region present in SUP35, are highlighted. Change of protein stability of the GSPT1–SUP35 fusion proteins in response to CC-885 treatment is shown in Extended Data Fig. 5d. **b**, Immunoblot analysis of 293T HEK cells stably producing HA-tagged GSPT1 wild-type and variants. Dotted line marks where irrelevant lanes were removed. Detailed domain mapping experiments are shown in Extended Data Fig. 5. **c**, Effect of CC-885 on proliferation of 293T parental cells or stably expressing GSPT1 variants.

sequence homology with human GSPT1, but did not show any changes in protein level in response to CC-885 when stably expressed in 293T HEK cells (Extended Data Fig. 5). Through a series of domain-swap experiments, we identified that conversion of GSPT1 glycine 575 into the corresponding asparagine of SUP35 prevented both the CC-885-induced degradation of GSPT1 and the *in vitro* binding to cereblon (Fig. 2a, Extended Data Fig. 5). We then tested CC-885 in 293T HEK cells stably expressing the CC-885-sensitive or -resistant GSPT1 variants. Overexpression of a resistant variant GSPT1 Δ (1–138)/(G575N) completely abrogated the CC-885-induced anti-proliferation (Fig. 2b, c), whereas overexpression of a CC-885-sensitive variant GSPT1 Δ (1–138) only conferred partial protection. Similar results were obtained in AML cell lines OCI-AML2 and MOLM-13 (Extended Data Fig. 6a, b). To determine whether GSPT1 depletion is sufficient to block cell proliferation, we used lentiviral short hairpin RNA (shRNA) vectors to silence the expression of endogenous GSPT1 in 293T HEK, OCI-AML2, and MOLM-13 cells and found a marked reduction of GSPT1 protein level correlated with decreased cell fitness as compared to parental cells or cells expressing a control shRNA (Fig. 2d, e and Extended Data Fig. 6c–e). Together, these experiments indicate that CC-885 cytotoxicity largely, if not completely, depends upon loss of GSPT1. This is consistent with a previously published report showing that mutation of SUP35 prevented the G1 to S phase transition during cell cycle progression in *Saccharomyces cerevisiae*²⁵. Similarly, depletion of GSPT1 led to cell cycle arrest at G1 in the human colorectal cancer cell line HCT116 (ref. 26).

Structure of cereblon with DDB1, CC-885 and GSPT1

To explore the molecular requirements for ligand-directed substrate recruitment to cereblon, we determined the crystal structure of full-length human DDB1 bound to human cereblon (amino acids 40–442), CC-885, and domains 2 and 3 of human GSPT1 (amino acids 437–633) to 3.6 Å resolution (Fig. 3a, full data collection and refinement statistics given in Extended Data Table 1). In the crystal structure, GSPT1 domain 3 docks against cereblon at the site of CC-885 binding with

Result is representative of three biological replicates. Data are presented as mean \pm s.d., $n = 3$ technical replicates. **d**, **e**, Cell growth curve of OCI-AML2 parental cells or cells transduced with lentiviral vectors expressing a control shRNA (shCNTL) or GSPT1 specific shRNA (shGSPT1-1 to 4). Immunoblot analysis of whole-cell extracts (**d**) was carried out at day 4 after transduction with lentiviral vectors. Cell growth (**e**) was quantified with CellTiter-Glo (CTG) at day 4, 6, and 8 after transduction. Results are representative of three biological replicates. Data are mean \pm s.d. $n = 3$ technical replicates. For gel source data, see Supplementary Information Fig. 1.

direct interactions with both CC-885 and the proximal cereblon surface (Fig. 3a). GSPT1 domain 2 does not directly contact either CC-885 or cereblon. Cereblon is bound to DDB1 in a manner consistent with the previously determined structures of cereblon–DDB1 (refs 10, 11). CC-885 is bound in the tri-Trp pocket of cereblon, where the glutarimide ring makes 3 hydrogen bonds with cereblon, two to the backbone of residues W380 and H378, and one to the side chain of H378 (Fig. 3b). The isoindolinone ring of CC-885 (Fig. 1a) is presented on the surface of cereblon in a manner similar to lenalidomide, and interacts with both cereblon and GSPT1, including a hydrogen bond from the isoindolinone carbonyl oxygen to the side chain of N351 of cereblon. The chemical structure of CC-885 is extended compared to lenalidomide or pomalidomide (Fig. 1a), allowing further interactions with both cereblon and GSPT1. The urea moiety of CC-885 is positioned between cereblon residues E377 and H353, with hydrogen bonds to both side chains. The terminal methyl-chloro-phenyl ring is positioned proximal to the β -sheet core of GSPT1 domain 3.

CC-885 is critical for the formation of the complex as demonstrated by surface plasmon resonance experiments showing that GSPT1 binds to cereblon–DDB1 with a dissociation constant (K_D) of ~ 350 nM in the presence of CC-885 (Extended Data Fig. 7). No significant interaction was observed in the absence of CC-885. Similarly, negative stain electron microscopy confirmed that CC-885 was critical for cereblon–GSPT1 complex formation, as no cereblon–GSPT1 particles were observed in the absence of CC-885 (Extended Data Fig. 8).

GSPT1 binds a ‘hotspot’ on the cereblon–CC-885 surface

It was previously hypothesized that immunomodulatory drug binding forms an interaction hotspot on the cereblon surface by placement of a planar hydrophobic group (the ligand phthalimide or isoindolinone ring, Fig. 1a) amongst unsatisfied hydrogen bonds from both ligand and the protein surface¹⁰. In this way, ligands with low molecular weight would be able to achieve marked alterations in binding potency by promoting contributions towards substrate binding from the cereblon surface. The structure of cereblon–DDB1–CC-885–GSPT1 confirms that GSPT1 interacts directly with CC-885 and also forms

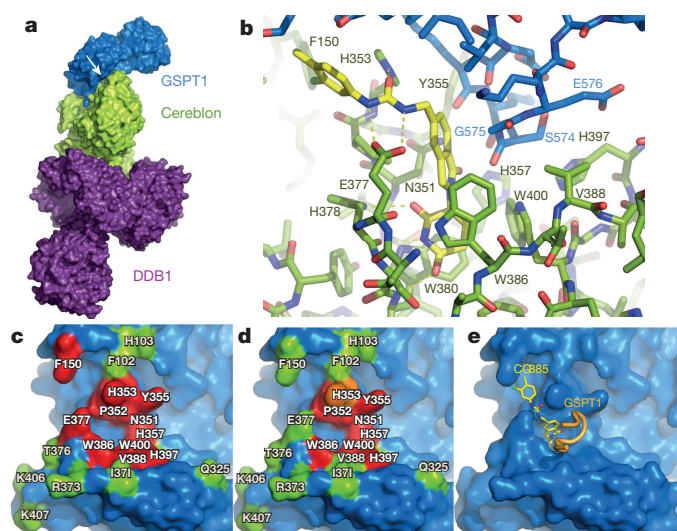


Figure 3 | Crystal structure of the cereblon-DDB1-GSPT1-CC-885 complex. **a**, Surface representation with DDB1 shown in purple, cereblon in green and GSPT1 in blue. Position of CC-885 is indicated by an arrow. **b**, Details of the interface between cereblon and GSPT1 with CC-885 shown as sticks with carbons in yellow. Hydrogen bonds between CC-885 and cereblon are shown as yellow dotted lines. **c–e**, The effect of cereblon surface mutagenesis on substrate binding (see also Extended Data Fig. 9). Red indicates a strong reduction in substrate binding following mutation to alanine, orange indicates some reduction, and green indicates no effect. Blue surface areas were not tested. **c**, The effect of the mutations on GSPT1 recruitment; **d**, the mutational results for Ikaros binding; **e**, the relative positions of CC-885 (yellow sticks) and the GSPT1 cereblon-interacting motif (orange).

direct protein–protein interactions, with cereblon residues N351, H357, and W400 all contributing direct hydrogen bonds to GSPT1. Further van der Waals interactions are provided from the side chains of cereblon residues Y355, H397, and V388.

Co-immunoprecipitation of GSPT1 with cereblon surface mutants reveals that residues including N351, H357, V388 and H397 are critical to recruiting GSPT1 to cereblon in a manner dependent upon CC-885 (Fig. 3c, Extended Data Fig. 9). Each of the three side chains that hydrogen bond to GSPT1 (N351, H357 and W400) is shown to be critical for GSPT1 recruitment. When mutation sensitivity is mapped onto the surface of cereblon by colour, a patch of sensitivity is revealed that corresponds to the interaction site visible in the crystal structure (Fig. 3c, e).

To evaluate whether this patch of the cereblon surface is similarly purposed for the recruitment of other substrates, the same panel of cereblon mutants was tested against Ikaros (Fig. 3d, Extended Data Fig. 9). As shown in Fig. 3c, d, the critical interaction site is similar between GSPT1 and Ikaros. Interestingly, all three residues that contribute hydrogen bonds to GSPT1 are also required for Ikaros recruitment, suggesting a common mechanism of substrate recruitment. Residues E377 and F150, however, are uniquely required for GSPT1 recruitment. F150 is proximal to CC-885 and contacts domain 3 of GSPT1 in a manner that is dependent on the GSPT1 tertiary structure; it is therefore probably different from other substrates that do not share a common fold. E377 makes direct interactions with CC-885, and mutation may therefore affect ligand binding and conformation. An E377V polymorphism is found in mouse and rats and underlies resistance to CC-885 treatment in rodents (Extended Data Fig. 9 e and Supplementary Discussion). A further difference in cereblon–substrate interactions is observed at V388: GSPT1 binding is lost in the V388A mutant, but some binding is retained in the V388I mutant, which is also a rodent polymorphism. In contrast, Ikaros binding is abolished by the V388I mutation, consistent with previous work⁵; however, the V388A mutation has no significant effect on Ikaros binding. These differences

may be due to subtle differences in residue side chain bulk or substrate conformation at the interaction site. Lenalidomide and pomalidomide do not cause GSPT1 binding and degradation (Fig. 1b and Extended Data Fig. 2). These molecules lack the urea and chloro-methyl-phenyl moieties of CC-885 (Fig. 1a) that probably contribute to the affinity of the complex through interactions with domain 3 of GSPT1.

GSPT1 features required for recruitment to cereblon

The GSPT1 motif that mediates interactions with cereblon–CC-885 is a solvent-exposed region composed of residues 569–578, including a small anti-parallel β -sheet forming a β -hairpin with an α -turn from residues 571–575 (Fig. 4a). Despite the degree of solvent exposure, the turn is well-ordered in both the previous GSPT1 crystal structure¹⁷ and in the complex with cereblon. The conformational stability of the turn is probably due to intramolecular bonding, with an ASX-motif and an ST-turn mediated by D571 and S574, respectively. These intramolecular motifs might be expected to rigidify and stabilize the conformation of the turn before cereblon binding (Fig. 4a). The key hydrogen bonds formed with cereblon are all mediated by backbone carbonyl oxygen atoms, which are displayed in an array at the end of the turn. Carbonyl oxygen atoms from GSPT1 K572, K573, and S574 accept hydrogen bonds from cereblon residues N351, H357, and W400, respectively (Fig. 4a). The GSPT1 turn lies on top of the isoindolinone ring of CC-885 making hydrophobic and van der Waals interactions. A further key interaction occurs between G575 and the isoindolinone moiety. G575 is of particular importance in determining substrate recruitment, as it not only contributes to binding interactions, but also the close sterics at this position indicate that no other residue would be tolerated at this position. Therefore, the degon is not the linear peptide sequence, but rather the geometric arrangement of three backbone hydrogen bond acceptors at the apex of a turn (positions i , $i+1$, and $i+2$), with a glycine residue at a key position ($i+3$).

The structural interpretation is supported by site-directed mutagenesis of the GSPT1 turn, which indicates that G575A loses CC-885-dependent cereblon binding (Fig. 4c). It should be noted that the backbone torsion angles of G575 occupy a part of the Ramachandran plot only accessible to glycine residues and mutagenesis may alter both side chain decoration and backbone position. V570 exhibits van der Waals interactions with the methylene–urea linker, and accordingly, the V570A mutant displays reduced cereblon binding (Fig. 4a, c). Consistent with their expected role in stabilizing the turn conformation, both D571A and S574A mutants lose cereblon binding (Fig. 4c). However, the side chains of residues K572, K573 and E576 do not affect GSPT1 binding to cereblon, consistent with the crystal structure that indicates these residues have minimal side chain interactions with cereblon. The fact that key bonds are formed from the peptide backbone of GSPT1, rather than from side chain residues, presents the intriguing possibility that there may be considerable sequence tolerance in substrate recruitment, as long as the backbone remains in the appropriate conformation.

Ikaros binds via a similar structural motif

GSPT1 contains no obvious structural or sequence homology to the previously identified cereblon substrates such as the protein kinase CK1 α or the zinc finger transcription factor Ikaros. However, the similar pattern of sensitivity to mutations on the surface of cereblon (Fig. 3c, d) suggests that a similar anchoring motif may be present in Ikaros. On the basis of the observation that the GSPT1 interactions with cereblon were mediated by the protein backbone, we constructed a homology model of Ikaros and examined this for a similar surface turn. The homology model of Ikaros indicated that a zinc finger domain is indeed capable of presenting a turn that can make the three key hydrogen bonds with a glycine residue at the critical position on the solvent-accessible surface of the domain (Fig. 4b). Consistent with this prediction, a minimal domain of Ikaros containing only this single predicted zinc finger domain (amino acids 140–168),

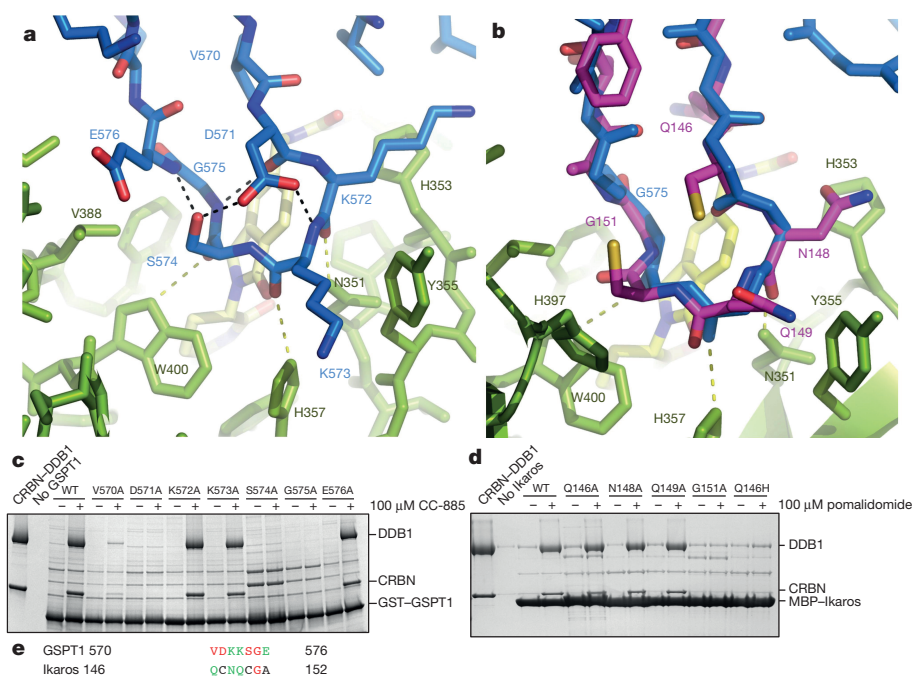


Figure 4 | Substrate proteins bind cereblon via a surface turn. **a**, The anchoring motif of GSPT1 (blue) bound to cereblon (green) and CC-885 (yellow). **b**, A model of Ikaros (magenta) superimposed onto the backbone of GSPT1 (blue). **c**, Coomassie-stained gels of cereblon–DDB1 pull-down by GST–GSPT1 (amino acids 526–633), with alanine mutations of the GSPT1 anchoring motif. **d**, Cereblon–DDB1 pull-down with MBP–Ikaros (amino acids 140–168) with alanine mutations of the proposed Ikaros

anchoring motif. This experiment is representative of three independent replicates (**c**, **d**). **e**, Sequence alignment of the GSPT1-anchoring motif with the putative region from Ikaros, with residues shown to be critical for binding in red, and those where an alanine mutation was tolerated in green. The critical glycine residue is shown in bold. Ikaros C147 and C150 were not mutated as they coordinate a structural zinc ion (not shown). For gel source data, see Supplementary Information Fig. 1.

is sufficient to pull down cereblon–DDB1 in the presence of pomalidomide (Fig. 4d).

Crucially, there is no sequence homology between Ikaros and GSPT1 in the anchoring motif regions as modelled, except for the glycine residue at position i+3 (Fig. 4e). The implications of this model were tested by co-immunoprecipitation of cereblon with mutations targeting each residue of the putative Ikaros anchoring motif. Mutation of the key glycine residue (G151) to alanine disrupts cereblon interactions (Fig. 4d). It was previously demonstrated that the Q146H mutation causes a reduction of cereblon binding^{15,16}. The putative anchoring motif modelled in Ikaros positions the critical residue, Q146, in a position oriented towards both bound ligand and the cereblon surface. Mutation of the putative anchoring motif in Ikaros confirms that Q146H causes a reduction in cereblon binding (Fig. 4d, Extended Data Fig. 2a). In contrast, the Q146A mutation is similar to wild type, suggesting that steric or electronic hindrance of a histidine residue at this position, rather than loss of key interactions from the glutamine, is responsible for the loss of cereblon binding in the Q146H mutants. N148A and Q149A mutations in the proposed anchoring motif of Ikaros do not affect cereblon binding, in a manner analogous to the equivalent positions K572 and K573 in GSPT1 (Fig. 4c, d). C147 and C150 are responsible for coordination of a structural zinc ion, and were therefore not examined by mutagenesis out of concern for disruption of the zinc finger domain structure. Structural studies on Ikaros family members in complex with cereblon will be necessary to identify the full atomic details of the interactions.

Discussion

In this work, we have extended the clinical scope of cereblon-modulating activities with the identification of a novel ligand, CC-885, which is potently anti-proliferative in cancer cell lines. CC-885 directs cereblon to target GSPT1, a substrate that is unrelated in fold and function to previously described substrates. Targeting of GSPT1

provides expanded clinical potential for cereblon modulators, with potent effects seen in both AML cell lines and in patient-derived tumour cells. Furthermore, by targeting what may have been considered an otherwise undruggable protein, the utility of directing cereblon for targeting diverse protein substrates posits a wide range of new therapeutic opportunities.

In contrast to linker-based approaches exploiting cereblon^{27–29}, CC-885 creates an interaction hotspot on the cereblon surface for direct protein–protein interactions with the substrate, underlying the gain-of-function ‘chemo-neomorphic’ activity. Although this may constrain the potential substrate range, the contribution from the cereblon surface supports the function of small drug-like molecules without the pharmacokinetic challenges that may accompany ligase modulators incorporating large flexible linkers.

Our structural work shows that cereblon recruits GSPT1 through the CC-885-induced recognition of a surface turn. A key to the steric compatibility is the presence of a glycine at the precise position within the turn. Significantly, except the glycine residue, the interactions are composed of hydrogen bonds from the backbone of GSPT1, indicating that the degron in this instance will not be revealed by primary sequence, but instead by the geometry and sterics. Our mutational and modelling work on Ikaros raises the intriguing possibility of a common structural degron occurring amongst diverse substrate proteins. Furthermore, after submission of this manuscript, a crystal structure of cereblon in complex with lenalidomide and CK1 α was published that exhibits a binding mode consistent with our degron³⁰. It may therefore be possible to rationally identify candidate neosubstrates by searching for the cereblon response element across the proteome. Such candidate substrates may then form targets for exploitation by the next generation of cereblon modulators.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 27 November 2015; accepted 31 May 2016.

Published online 22 June 2016.

1. Ito, T. *et al.* Identification of a primary target of thalidomide teratogenicity. *Science* **327**, 1345–1350 (2010).
2. Dimopoulos, M. A., Richardson, P. G., Moreau, P. & Anderson, K. C. Current treatment landscape for relapsed and/or refractory multiple myeloma. *Nat. Rev. Clin. Oncol.* **12**, 42–54 (2015).
3. Lopez-Girona, A. *et al.* Cereblon is a direct protein target for immunomodulatory and antiproliferative activities of lenalidomide and pomalidomide. *Leukemia* **26**, 2326–2335 (2012).
4. Zhu, Y. X. *et al.* Cereblon expression is required for the antimyeloma activity of lenalidomide and pomalidomide. *Blood* **118**, 4771–4779 (2011).
5. Krönke, J. *et al.* Lenalidomide induces ubiquitination and degradation of CK1 α in del(5q) MDS. *Nature* **523**, 183–188 (2015).
6. Angers, S. *et al.* Molecular architecture and assembly of the DDB1–CUL4A ubiquitin ligase machinery. *Nature* **443**, 590–593 (2006).
7. Higa, L. A. *et al.* CUL4–DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. *Nat. Cell Biol.* **8**, 1277–1283 (2006).
8. Jin, J., Arias, E. E., Chen, J., Harper, J. W. & Walter, J. C. A family of diverse Cul4–Ddb1-interacting proteins includes Cdt1, which is required for S phase destruction of the replication factor Cdt1. *Mol. Cell* **23**, 709–721 (2006).
9. He, Y. J., McCall, C. M., Hu, J., Zeng, Y. & Xiong, Y. DDB1 functions as a linker to recruit receptor WD40 proteins to CUL4–ROC1 ubiquitin ligases. *Genes Dev.* **20**, 2949–2954 (2006).
10. Chamberlain, P. P. *et al.* Structure of the human Cereblon–DDB1–lenalidomide complex reveals basis for responsiveness to thalidomide analogs. *Nat. Struct. Mol. Biol.* **21**, 803–809 (2014).
11. Fischer, E. S. *et al.* Structure of the DDB1–CRBN E3 ubiquitin ligase in complex with thalidomide. *Nature* **512**, 49–53 (2014). 10.1038/nature13527.
12. John, L. B. & Ward, A. C. The Ikaros gene family: transcriptional regulators of hematopoiesis and immunity. *Mol. Immunol.* **48**, 1272–1278 (2011).
13. Dijon, M. *et al.* The role of Ikaros in human erythroid differentiation. *Blood* **111**, 1138–1146 (2008).
14. Gandhi, A. K. *et al.* Immunomodulatory agents lenalidomide and pomalidomide co-stimulate T cells by inducing degradation of T cell repressors Ikaros and Aiolos via modulation of the E3 ubiquitin ligase complex CRL4(CRBN). *Br. J. Haematol.* **164**, 811–821 (2014).
15. Krönke, J. *et al.* Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science* **343**, 301–305 (2014).
16. Lu, G. *et al.* The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. *Science* **343**, 305–309 (2014).
17. Cheng, Z. *et al.* Structural insights into eRF3 and stop codon recognition by eRF1. *Genes Dev.* **23**, 1106–1118 (2009).
18. Preis, A. *et al.* Cryoelectron microscopic structures of eukaryotic translation termination complexes containing eRF1–eRF3 or eRF1–ABCE1. *Cell Reports* **8**, 59–65 (2014).
19. Zhouravleva, G. *et al.* Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. *EMBO J.* **14**, 4065–4072 (1995).
20. Hon, W. C. *et al.* Structural basis for the recognition of hydroxyproline in HIF-1 α by pVHL. *Nature* **417**, 975–978 (2002).
21. Min, J. H. *et al.* Structure of an HIF-1 α –pVHL complex: hydroxyproline recognition in signaling. *Science* **296**, 1886–1889 (2002).
22. Glotzer, M., Murray, A. W. & Kirschner, M. W. Cyclin is degraded by the ubiquitin pathway. *Nature* **349**, 132–138 (1991).
23. Sheard, L. B. *et al.* Jasmonate perception by inositol-phosphate-potentialized COI1–JAZ co-receptor. *Nature* **468**, 400–405 (2010).
24. Tan, X. *et al.* Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* **446**, 640–645 (2007).
25. Kikuchi, Y., Shimatake, H. & Kikuchi, A. A yeast gene required for the G1-to-S transition encodes a protein containing an A-kinase target site and GTPase domain. *EMBO J.* **7**, 1175–1182 (1988).
26. Chauvin, C., Salhi, S. & Jean-Jean, O. Human eukaryotic release factor 3a depletion causes cell cycle arrest at G1 phase through inhibition of the mTOR pathway. *Mol. Cell. Biol.* **27**, 5619–5629 (2007).
27. Deshaies, R. J. Protein degradation: prime time for PROTACs. *Nat. Chem. Biol.* **11**, 634–635 (2015).
28. Winter, G. E. *et al.* Drug development. Phthalimide conjugation as a strategy for *in vivo* target protein degradation. *Science* **348**, 1376–1381 (2015).
29. Lu, J. *et al.* Hijacking the E3 ubiquitin ligase cereblon to efficiently target BRD4. *Chem. Biol.* **22**, 755–763 (2015).
30. Petzold, G., Fischer, E. S. & Thomä, N. H. Structural basis of lenalidomide-induced CK1 α degradation by the CRL4(CRBN) ubiquitin ligase. *Nature* **532**, 127–130 (2016).

Supplementary Information is available in the online version of the paper.

Acknowledgements Thanks to G. Reyes, C. Havens, P. Jackson and H. Hadjivassiliou for discussions relating to this manuscript, and P. Jackson, J. Hansen, M. Correa, B. Fahr, M. Abbasian, E. Ambing, E. Rychak, D. Mendy and K. Hughes for technical assistance. We thank K. Motamedchaboki (Proteomics Core, Sanford Burnham Prebys Medical Discovery Institute) for mass spectrometry-based proteomic analysis. Thanks to J. Ballesteros and P. Hernandez at Viviva for patient sample testing. Parts of this work were conducted at the Advanced Light Source. The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health, National Institute of General Medical Sciences, and the Howard Hughes Medical Institute. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under contract No. DE-AC02-05CH11231. This work was also supported by Grant-in-Aid for Scientific Research on Innovative Areas “Chemical Biology of Natural Products” from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) (23102002 to H.H.), by JST, PRESTO (T.I.) and by Grant-in-Aid for Young Scientists (B) from MEXT (26750374 to T.I.).

Author Contributions M.E.M., P.P.C., B.P.W.F., J.H., T.I., G.C., and M.R. performed biochemical and crystallographic structural studies. G.L., C.-C.L., K.M., N.-Y.W., D.N., T.T., S.G., and S.X. performed molecular and cellular biology experiments. G.C.L., L.N., M.E.M., and P.P.C. performed electron microscopy studies. A.L.R. performed chemical synthesis. M.E.M., G.L., A.L.-G., J.C., T.I., H.H., T.O.D., B.C., and P.P.C. planned the work, and all authors contributed to the manuscript.

Author Information Coordinates for the structure of cereblon–DDB1–CC-885–GSPT1 have been deposited in the Protein Data Bank with the accession code 5HXB. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.P.C. (pchamberlain@celgene.com) or G.L. (glu@celgene.com).

METHODS

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Protein expression and purification. ZZ-domain-6×His-thrombin-tagged human cereblon (amino acids 40–442) and full-length human DDB1 were co-expressed in SF9 insect cells in ESF921 medium (Expression Systems), in the presence of 50 μM zinc acetate. Cells were resuspended in buffer containing 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10 mM imidazole, 10% glycerol, 2 mM TCEP, 1× Protease Inhibitor Cocktail (San Diego Bioscience), and 40,000 U Benzonase (Novagen), and sonicated for 30 s. Lysate was clarified by high speed spin at 108,800g for 30 min, and clarified lysate was incubated with Ni-NTA affinity resin (Qiagen) for 1 h. Complex was eluted with buffer containing 500 mM imidazole, and the ZZ-domain-6×His tag removed by thrombin cleavage (Enzyme Research) overnight, combined with dialysis in 10 mM imidazole buffer. Cleaved eluate was incubated with Ni-NTA affinity resin (Qiagen), and the flow-through diluted to 200 mM NaCl for further purification over an ANX HiTrap ion exchange column (GE Healthcare). The ANX column was washed with ten column volumes of 50 mM Tris-HCl (pH 7.5), 200 mM NaCl, 3 mM TCEP, followed by ten column volumes of 50 mM Bis-Tris (pH 6.0), 200 mM NaCl, 3 mM TCEP, and the cereblon–DDB1 peak eluted at 210 mM NaCl. This peak was collected and further purified by size-exclusion chromatography using a Sephacryl S-400 16/60 column (GE Healthcare) in buffer containing 10 mM HEPES (pH 7.0), 240 mM NaCl, and 3 mM TCEP. The cereblon–DDB1 complex was concentrated to 30 mg ml^{−1} for crystallization trials.

GSPT1 domains 2 and 3 (amino acids 437–633) were expressed as an MBP-fusion in *Escherichia coli* BL21 (DE3) Star cells (Life Technologies) using 2XYT media (Teknova). Cells were induced at optical density OD₆₀₀ 0.6 for 18 h at 16°C. Cells were pelleted, resuspended in buffer containing 200 mM NaCl, 50 mM Tris (pH 7.5), 1 mM TCEP, 10% glycerol, 0.01 mg ml^{−1} lysozyme (Sigma), 10,000 U benzonase (Novagen), and 1× protease inhibitor cocktail (San Diego Bioscience). Resuspended cells were frozen, thawed for purification, and sonicated for 30 s before high speed spin at 108,800g for 30 min. Clarified lysate was incubated with amylose resin (NEB) at 4°C for 1 h before beads were washed with 500 ml wash buffer containing 200 mM NaCl, 50 mM Tris (pH 7.5), 1 mM TCEP, 10% glycerol. Protein was eluted with buffer containing 10 mM maltose, and the MBP tag was removed by overnight cleavage with thrombin (Enzyme Research). Cleaved GSPT1 was diluted into buffer containing 90 mM NaCl and further separated over a Heparin HiTrap column (GE healthcare). The GSPT1 peak eluting at 100 mM NaCl was collected, concentrated, and further separated by size exclusion chromatography using a Superdex 75 16/600 column (GE Healthcare) in buffer containing 10 mM HEPES pH 7, 240 mM NaCl, and 3 mM TCEP. The peak containing GSPT1 domains 2 and 3 eluted at 75 ml and was concentrated to 16 mg ml^{−1} for crystallization trials.

MBP-Ikaros 140–168 and mutants were expressed in *E. coli* BL21 (DE3) Star cells (Life Technologies) using 2XYT media (Teknova). Cells were induced at OD₆₀₀ 0.6 for 18 h at 16°C. Cells were pelleted, resuspended in buffer containing 200 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, 10% glycerol, 150 μM zinc acetate, 0.01 mg ml^{−1} lysozyme (Sigma), 40,000 U benzonase (Novagen), and 1× protease inhibitor cocktail (San Diego Bioscience). Resuspended cells were frozen, thawed for purification, and sonicated for 30 s before high-speed spin at 108,800g for 30 min. Clarified lysate was incubated with amylose resin (NEB) at 4°C for 1 h before beads were washed. Protein was eluted with buffer containing 200 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, 10% glycerol, 150 μM zinc acetate, and 10 mM maltose. Eluate was concentrated and further purified by size exclusion chromatography over a Superdex 200 16/600 column (GE Healthcare) in buffer containing 200 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, 10% glycerol, and 150 μM zinc acetate.

GST–GSPT1 domain 3 (amino acids 526–633) and mutants were expressed in *E. coli* BL21 (DE3) Star cells (Life Technologies) using 2XYT media (Teknova). Cells were induced at OD₆₀₀ 0.6 for 18 h at 16°C. Cells were pelleted, resuspended in buffer containing 200 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, 10% glycerol, 0.01 mg ml^{−1} lysozyme (Sigma), 20,000 U benzonase (Novagen), and 1× protease inhibitor cocktail (San Diego Bioscience). Resuspended cells were frozen, thawed for purification, and sonicated for 30 s before high-speed spin at 108,800g for 30 min. Clarified lysate was incubated with Glutathione Sepharose 4 Fast Flow (GE Healthcare) at 4°C for 2 h before beads were washed. Protein was eluted with buffer containing 200 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, 10% glycerol, and 20 mM reduced glutathione. Eluate was concentrated and further purified by size-exclusion chromatography over a Superdex 200 16/600 column (GE Healthcare) in buffer containing 200 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, and 10% glycerol.

6×His–eRF1 (human, full length) was expressed in *E. coli* BL21 (DE3) Star cells (Life Technologies) using 2XYT media (Teknova). Cells were induced at OD₆₀₀ 0.6 for 18 h at 16°C. Cells were pelleted, resuspended in buffer containing 150 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, 10% glycerol, 0.01 mg ml^{−1} lysozyme (Sigma), 20,000 U benzonase (Novagen), and 1× protease inhibitor cocktail (San Diego Bioscience). Resuspended cells were frozen, thawed for purification, and sonicated for 30 s before high speed spin at 108,800g for 30 min. Clarified lysate was incubated with Ni-NTA affinity resin (Qiagen) at 4°C for 1 h before beads were washed. Protein was eluted with buffer containing 150 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, 10% glycerol, and 500 mM imidazole. Eluate was concentrated and further purified by size-exclusion chromatography over a Superdex 200 16/600 column (GE Healthcare) in buffer containing 150 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, and 10% glycerol.

6×His–DDB1 (human, full length) was expressed in SF9 insect cells in ESF921 medium (Expression Systems). Cells were resuspended in buffer containing 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10 mM imidazole, 10% glycerol, 2 mM TCEP, 1× Protease Inhibitor Cocktail (San Diego Bioscience), and 40,000 U Benzonase (Novagen), and sonicated for 30 s. Lysate was clarified by high-speed spin at 108,800g for 30 min, and clarified lysate was incubated with Ni-NTA affinity resin (Qiagen) for 1 h. Complex was eluted with buffer containing 500 mM imidazole. Eluate was collected and further purified by size-exclusion chromatography over a Superdex 200 16/600 column (GE Healthcare) in buffer containing 150 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, and 10% glycerol.

The V protein of simian virus 5 (SV5-V) (full length) was expressed in *E. coli* BL21 (DE3) Star cells (Life Technologies) using 2XYT media (Teknova) as a GST-fusion protein. Cells were induced at OD₆₀₀ 0.6 for 18 h at 16°C. Cells were pelleted, resuspended in buffer containing 150 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, 10% glycerol, 0.01 mg ml^{−1} lysozyme (Sigma), 20,000 U benzonase (Novagen), and 1× protease inhibitor cocktail (San Diego Bioscience). Resuspended cells were frozen, thawed for purification, and sonicated for 30 s before high speed spin at 108,800g for 30 min. Clarified lysate was incubated with Glutathione Sepharose 4 Fast Flow (GE Healthcare) at 4°C for 2 h before beads were washed and protein eluted with buffer containing 150 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, 10% glycerol, and 20 mM reduced glutathione. One half of the eluate was incubated with thrombin to remove the GST tag, while the other half retained the tag. Tagged and un-tagged proteins were further purified by size-exclusion chromatography over a Superdex 200 16/600 column (GE Healthcare) in buffer containing 150 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, and 10% glycerol.

SV5-V–DDB1 complex for the *in vitro* ubiquitination assay was generated by incubating stoichiometric amounts of purified, un-tagged SV5-V and purified 6×His–DDB1 overnight at 4°C before purification of the complex by separation over a Superdex 200 16/600 column (GE Healthcare) in buffer containing 150 mM NaCl, 50 mM Tris (pH 7.5), 3 mM TCEP, and 10% glycerol.

Negative stain electron microscopy analysis of GSPT1 binding to cereblon–DDB1–CC-885. To prepare grids for negative stain analysis of isolated cereblon–DDB1 or cereblon–DDB1 bound to CC-885 and GST–GSPT1 domains 2 and 3 (amino acids 437–633), pre-incubated protein complex samples were rapidly diluted from 30 μM to 0.3 μM in a buffer containing 50 mM Tris (pH 7.5), 150 mM NaCl, and 1 mM TCEP. A 4 μl aliquot was applied to freshly plasma-cleaned 400 mesh Cu–Rh maxtaform grids (Electron Microscopy Sciences) that had been coated with a thin layer of carbon. After incubating for 1 min, excess protein was wicked off with Whatman filter paper and the grid was immediately inverted and placed on a 50 μl droplet of 2% (w/v) uranyl formate solution. After 30 s, excess stain was wicked off from the grid by touching the edge with filter paper. This staining step was repeated three times for thorough embedding of the sample, and the grids were air dried on the edge of a fume hood for constant air flow after the last blotting step. Samples were imaged on a Tecnai Spirit operating at 120 keV. Data were collected using the Legikon automated image acquisition software (PMID 15890530), and analysed using the Appion processing package (PMID 9263523).

Crystallization and structure determination. Crystallization of the complex was achieved by sitting-drop vapour diffusion. Cereblon–DDB1 and GSPT1 were mixed together to equimolar stoichiometry at a final concentration of 150 μM. The solution of cereblon–DDB1–GSPT1 in the presence of 500 μM CC-885 was mixed 1:1 with, and subsequently equilibrated against, a mother liquor solution of 200 mM sodium citrate, Tris (pH 8.4–8.6), 17–20% PEG 3350 and incubated at 9°C. Crystals were cryoprotected in the reservoir solution supplemented with 20% ethylene glycol and cooled under liquid nitrogen. Data was collected from a single crystal at the Advanced Light Source, beamline 5.0.2. The structure of human cereblon–DDB1–GSPT1–CC-885 was solved by molecular replacement using Phaser³¹, with human cereblon–DDB1 (PDB code 4TZ4) and GSPT1 (PDB code 3E1Y) as search models. Subsequent manual model building using Coot and refinement were performed using Refmac5 with non-crystallographic symmetry and external structure restraints^{32,33}. Crystallographic statistics are summarized

in Extended Data Table 1, sample electron density is shown in Extended Data Fig. 7d. Ikaros was modelled on the basis of sequence homology to coordinates with the Protein Data Bank accession number 2II3.

Surface plasmon resonance binding assay. Binding kinetics were measured using a Biacore T200. Using the GST-capture kit (GE Healthcare, BR-1002-23), ~2000 relative units (RU) of anti-GST antibody was immobilized on a CM5 chip (GE Healthcare). GST-GSPT1 domains 2 and 3 ($0.07 \mu\text{g ml}^{-1}$) or GST ($0.15 \mu\text{g ml}^{-1}$, reference channel) was then flowed over the chip at $10 \mu\text{l min}^{-1}$ for 2 min, resulting in a capture level of approximately 10 RU. A threefold dilution series of cereblon-DDB1 ($3 \mu\text{M}$ to 3 nM) was then flowed over both channels at a rate of $30 \mu\text{l min}^{-1}$ for 300 s contact time, followed by 900 s dissociation time. After each cycle of binding and dissociation, the surface was regenerated with 10 mM glycine (pH 2.1) for 120 s at $30 \mu\text{l min}^{-1}$, followed by 3 M MgCl_2 for 60 s at $30 \mu\text{l min}^{-1}$. All reagents (GST, GST-GSPT1 domain 2 and 3, cereblon-DDB1, and running buffer) were in 10 mM HEPES (pH 7.4), 150 mM NaCl, 0.05% v/v Surfactant P20 containing 0.02% DMSO and saturating levels of the relevant compound, glutarimide or CC-885 (CC-885 at $10 \mu\text{M}$, glutarimide at $100 \mu\text{M}$). The apparent on- and off-rate constants were globally fit with a 1:1 kinetic binding model to the sensograms (black lines) using the Biacore T200 kinetic analysis software package.

Pull-down binding assays of cereblon and substrates using purified recombinant components. Pull-down of cereblon-DDB1 with either GSPT1 or Ikaros was performed using the same protocol. Tagged substrate, either GST-GSPT1-domain 3 (amino acids 526–633) or MBP-Ikaros (amino acids 140–168), was expressed in bacteria and purified using affinity and size-exclusion chromatography. Substrate was then bound to magnetic glutathione (Pierce) or amylose (NEB) beads for 1 h. Beads were washed three times in pull-down buffer (50 mM Tris pH 7.5, 150 mM NaCl, 10% glycerol, 0.01% NP-40, and 1 mM TCEP, with the addition $150 \mu\text{M}$ zinc acetate for Ikaros) to remove excess unbound substrate, before the addition of DMSO or $100 \mu\text{M}$ compound and $50 \mu\text{M}$ cereblon-DDB1, purified from insect cells using affinity, ion exchange, and size-exclusion chromatography. After 1 h incubation with occasional gentle mixing at room temperature, unbound cereblon-DDB1 was removed and beads were washed three times in pull-down buffer, for a total of 30 s. Substrate and bound cereblon-DDB1 was eluted using either 20 mM reduced glutathione or 10 mM maltose, separated by SDS-PAGE, and Coomassie stained.

Tagged GST-GSPT1 domains 2 and 3 (amino acids 437–633) was expressed in bacteria and purified using affinity and size-exclusion chromatography. GST or GST-GSPT1 domains 2 and 3 was then bound to magnetic glutathione (Pierce) beads for 1 h. Beads were washed three times in pull-down buffer (50 mM Tris pH 7.5, 150 mM NaCl, 10% glycerol, 0.01% NP-40, and 1 mM TCEP) to remove excess unbound substrate, before the addition of DMSO or $100 \mu\text{M}$ compound and $50 \mu\text{M}$ cereblon-DDB1, purified from insect cells using affinity, ion exchange, and size-exclusion chromatography. After 1 h incubation with occasional gentle mixing at room temperature, unbound cereblon-DDB1 was removed and beads were washed three times in pull-down buffer, for a total of 30 s. Substrate and bound cereblon-DDB1 was eluted using either 20 mM reduced glutathione, separated by SDS-PAGE, and Coomassie stained.

Pull-down of eRF1 or cereblon-DDB1 with GSPT1 domains 2 and 3 and CC-885. Tagged GST-GSPT1 domains 2 and 3 (amino acids 437–633) was expressed in bacteria and purified using affinity and size exclusion chromatography. GST or GST-GSPT1 domains 2 and 3 was then bound to magnetic glutathione (Pierce) beads for 1 h. Beads were washed three times in pull-down buffer (50 mM Tris (pH 7.5), 150 mM NaCl, 10% glycerol, 0.01% NP-40, and 1 mM TCEP) to remove excess unbound substrate, before the addition of DMSO or $100 \mu\text{M}$ compound and $50 \mu\text{M}$ purified cereblon-DDB1 or $30 \mu\text{M}$ purified eRF1, \pm CC-885. After 1 h incubation with occasional gentle mixing at room temperature, unbound cereblon-DDB1 or eRF1 was removed and beads were washed three times in pull-down buffer, for a total of 30 s. Substrate and bound cereblon-DDB1 or eRF1 was eluted using either 20 mM reduced glutathione, separated by SDS-PAGE, and Coomassie stained.

Pull-down binding assays of SV5-V, DDB1, and CUL4-RBX1. To confirm that purified SV5-V binds to purified DDB1, and that this complex then binds to CUL4-RBX1, we performed two consecutive GST pull-down assays. GST-SV5-5 or GST alone was purified from bacteria using affinity and size exclusion chromatography. Purified GST-SV5-V or GST was then bound to magnetic glutathione beads (Pierce) for 1 h. Beads were washed three times in pull-down buffer (50 mM Tris (pH 7.5), 150 mM NaCl, 10% glycerol, 0.01% NP-40, and 1 mM TCEP) to remove excess GST-protein, before the addition of purified DDB1. After 1 h of incubation with occasional mixing at room temperature, unbound DDB1 was removed and beads were washed 3 times in pull-down buffer. Substrate and bound DDB1 were then eluted using either 20 mM reduced glutathione, separated by SDS-PAGE, and Coomassie stained. For subsequent binding of CUL4-RBX1, the same pull-down described above was used to generate GST-SV5-V-DDB1 on beads. However, after the beads were washed to remove excess DDB1, CUL4-RBX1

was added and incubated for 1 h at room temperature before beads were washed three times in pull-down buffer and bound protein eluted with 20 mM reduced glutathione, separated by SDS-PAGE, and Coomassie stained.

In vivo degradation of cereblon substrates. 293FT (Invitrogen) *CRBN*^{-/-} cells transiently transfected with plasmids expressing V5-tagged IKZF1, Flag-tagged Ikaros (IKZF3), Myc-tagged GSPT1, GFP and human or mouse cereblon variants. 36 h after transfection, cells were treated with DMSO, $10 \mu\text{M}$ lenalidomide, or $1 \mu\text{M}$ CC-885 for an additional 12 h. Cells were then washed with ice-cold $1 \times$ PBS twice, lysed in buffer A (50 mM Tris-HCl, 150 mM NaCl, 1% Triton-X 100, complete protease inhibitor tablet (Roche), phosphatase inhibitor tablet (Roche)). Whole-cell extracts were collected and subjected to immunoblot analysis.

Mammalian lysate co-immunoprecipitation. 293FT (Invitrogen) *CRBN*^{-/-} cells were transiently transfected with plasmids expressing Flag-tagged cereblon variants and HA-tagged GSPT1 (or HA-tagged IKZF1). After 2 days, cells were lysed in Buffer B (50 mM Tris (pH 7.4), 150 mM NaCl, 0.5% NP-40, 10% glycerol, $1 \times$ Complete Ultra protease inhibitor (Roche), and $1 \times$ PhosphoSTOP (Roche)). Whole-cell extracts were collected after centrifugation at $16,000g$ for 10 min. Whole-cell extracts containing cereblon and GSPT1 (or IKZF1) were then incubated with anti-HA affinity resin (Roche) in the presence of DMSO or $10 \mu\text{M}$ CC-885 (or $10 \mu\text{M}$ lenalidomide). For testing the interaction between cereblon variants and DDB1, whole-cell extracts expressing Flag-CRBN were incubated with anti-Flag affinity resin (Sigma). After overnight incubation at 4°C , anti-Flag and anti-HA resins were washed with Buffer B six times. Anti-Flag and anti-HA immunoprecipitates were then subjected to immunoblot analysis.

Cell culture and materials. Human embryonic kidney cell lines 293FT (Invitrogen), 293T (ATCC) and Lenti-X 293 (Clontech) were maintained in Dulbecco's Modified Eagle's medium (DMEM; Invitrogen) supplemented with 10% fetal bovine serum (FBS; Invitrogen), $1 \times$ sodium pyruvate (Invitrogen), $1 \times$ non-essential amino acids (Invitrogen), 100 U ml^{-1} penicillin (Invitrogen), and $100 \mu\text{g ml}^{-1}$ streptomycin (Invitrogen). Acute myeloid leukaemia cell lines KG-1, Kasumi-1, U937, MOLM-13, HL-60, and MV-4-11 were purchased from American Tissue Culture Collection (ATCC). NB-4, HNT-34, OCI-AML2, and OCI-AML3 cell lines were purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH. KG-1, Kasumi-1, U937, MOLM-13, NB-4, and HNT-34 cell lines were maintained in Roswell Park Memorial Institute (RPMI) 1640 tissue culture medium (Invitrogen) supplemented with 10% FBS, $1 \times$ sodium pyruvate, $1 \times$ non-essential amino acids, 100 U ml^{-1} penicillin, and $100 \mu\text{g ml}^{-1}$ streptomycin. HL-60 and MV-4-11 cell lines were maintained in Iscove's Modified Dulbecco's medium (IMDM; Invitrogen) supplemented with 10% FBS, $1 \times$ sodium pyruvate, $1 \times$ non-essential amino acids, 100 U ml^{-1} penicillin, and $100 \mu\text{g ml}^{-1}$ streptomycin. OCI-AML2 and OCI-AML3 cell lines were maintained in minimal essential medium (MEM; Invitrogen) supplemented with 10% FBS, $1 \times$ sodium pyruvate, $1 \times$ non-essential amino acid, 100 U ml^{-1} penicillin, and $100 \mu\text{g ml}^{-1}$ streptomycin. All cell lines were cultured at 37°C with 5% CO_2 in the relevant media mentioned above. All cell lines were confirmed to be mycoplasma-negative using the MycoAlert Mycoplasma Detection Kit (Lonza).

Cell proliferation assay. Human cancer cell lines cultured in the growth medium recommended by the vendor were seeded into black 384-well plates containing DMSO or test compounds. The seeding density for each cell line was optimized to allow the cell growth in the linear range during a 3-day culture period. To test the compound effect on cell proliferation in AML cell lines, 5,000 to 10,000 cells per well in $200 \mu\text{l}$ complete culture media were seeded into black 96-well plates containing DMSO or test compounds. After 48 or 72 h, cell proliferation was assessed using the CellTiter-Glo (CTG) Luminescent Cell Viability Assay (Promega) according to manufacturer's instructions. To test the effect of GSPT1 depletion on cell proliferation in AML cell lines, cells were infected with lentiviral shRNA vectors for 4 days, and cell proliferation was quantified using CTG every other day thereafter. Relative cell proliferation was normalized against day 4 cell growth values. To determine the effect of GSPT1 depletion on cell proliferation in 293T HEK cells, cells were infected with lentiviral shRNA vectors for 7 days, and then imaged using the EVOS FL Cell Imaging System (Thermo Fisher). The effect of CC-885 on AML patient-derived cells was performed by automated flow cytometry according to the methods described in ref. 34. Briefly, patient-derived bone marrow samples from adult patients over 18 years of age who were diagnosed with AML, were obtained from Hospital La Fe, Hospital General Universitario Gregorio Marañón and Hospital Lucus Augusti hospitals, all in Spain, following IRB approved protocols and signed informed consent. Bone marrow samples were extracted under sterile conditions and were received in the laboratory within 24 h of extraction where they were immediately processed. A small portion of the sample was separated for validation and selection of the monoclonal antibodies (mAb) to optimize the identification of leukaemic cells. The majority of the sample was diluted with culture media and plated into 96-well plates previously prepared with the test compounds. These plates were incubated for 24, 48 and 96 h and then

processed via the ExviTech platform. For this purpose, red cells were lysed and leukocytes were stained using the mAb previously selected and with Annexin-V to exclude apoptotic cells.

Antibodies. Rabbit anti-human CRBN65 mAb (Celgene, San Diego, CA)³, rabbit anti-DYKDDDDK (Flag) (9A3) mAb (#8146, Cell Signaling), mouse anti-Myc-Tag (9B11) mAb (#2276, Cell Signaling), mouse anti-GFP (B-2) mAb (#sc-9996, Santa Cruz), rabbit anti-human GSPT1 polyclonal antibody (pAb) (ab49878, Abcam), rabbit anti-human GSPT1 pAb (ab126090, Abcam), rabbit anti-GSPT1 pAb (10763-1-AP, Proteintech), rabbit anti-Ikaros (D10E5) mAb (#9034, Cell signaling), mouse anti-HA.11 (16B12) mAb (#901501, Biolegend), rabbit anti-DDB1 pAb (#5428, Cell Signaling), mouse anti-ubiquitin P4D1 mAb (#SC-8017, Santa Cruz), mouse anti-SV5-V (Santa Cruz, sc-58052), mouse anti-HIS mAb antibody (34660, Qiagen), mouse anti-actin mAb (#A5316, Sigma) and mouse anti-tubulin mAb (#T5201, Sigma) were used as primary antibodies. Goat anti-mouse 800 antibody (#926-32210, LICOR Biosciences), goat anti-rabbit 680 antibody (#926-68071, LI-COR Biosciences), goat anti-mouse 680 antibody (#926-68070, LI-COR Biosciences) and goat anti-rabbit 800 antibody (#926-32211, LI-COR Biosciences) were used as secondary antibodies.

Cellular immunoprecipitation. Cells were lysed in Buffer B (50 mM Tris (pH 7.4), 150 mM NaCl, 0.5% NP-40, one tablet of Complete ULTRA protease inhibitor cocktail (Roche), and one tablet of PhosSTOP phosphatase inhibitor cocktail (Roche)). Whole-cell lysates were clarified by centrifugation at top speed for 10 min. To identify cereblon-associated proteins, CC-885 (2 μ M) or DMSO were added into whole cell extracts of 293T HEK cells stably expressing Flag-HA-CRBN for 2 h. FH-CRBN complexes were then captured with anti-Flag beads (Sigma), eluted by Flag peptide, and separated on SDS-PAGE gel, followed by mass spectrometry or immunoblot analysis. To test the binding of cereblon to its substrates, 293FT HEK cells stably expressing HA-tagged GSPT1, IKZF1, IKZF1(Q146H) were lysed in Buffer B. HA-tagged recombinant proteins were then immunoprecipitated with anti-HA beads (Roche). After three washes with Buffer B, the anti-HA beads were mixed with cell lysates of 293FT HEK cells stably expressing shGSPT1-1 in the presence of DMSO, 10 μ M lenalidomide or 10 μ M CC-885 for 12 h. To test the binding of HA-tagged GSPT1 variants with cereblon and ERF1, whole-cell lysates of 293T HEK cells stably expressing GSPT1 variants were treated with DMSO or 20 μ M CC-885 for 2 h, followed by anti-HA immunoprecipitation for an additional 4 h. Beads were then washed with Buffer B six times, and bound proteins were eluted by boiling in LDS loading buffer and detected by immunoblot analysis.

Plasmids. Plasmids pDONR221 and pDONR223-IKZF3 were purchased from Invitrogen and assembled to generate pDONR223. Coding sequences of IKZF1, SUP35, GSPT1, cereblon and Flag-HA-cereblon were *in vitro* synthesized by Invitrogen and then cloned into pDONR223. Open reading frames encoding IKZF1(Q146H) and chimaeric fusion proteins of SUP35 and GSPT1 were generated via overlapping PCR, and then cloned into pDONR223. GSPT1 mutations and truncation deletions were introduced into pDONR223-GSPT1 via overlapping PCR. Cereblon mutants E377V and W386A in pDONR223 were generated via overlapping PCR.

DNA fragments of gateway-3 \times HA-mPGK (mouse phosphoglycerate kinase promoter)-GFP-P2A-Pur-WPRE, 3 \times HA-gateway-mPGK-Pur-WPRE and gateway-3 \times HA-mPGK-Pur-WPRE were *in vitro* synthesized by Invitrogen, and assembled into plenti6.2/v5-DEST (Invitrogen) to generate plenti-CMV-gateway-3 \times HA-mPGK-GFP-P2A-Pur, plenti-CMV-3 \times HA-gateway-mPGK-Pur and plenti-CMV-gateway-3 \times HA-mPGK-Pur, respectively. The Ubc promoter from plenti-6/Ubc/V5-DEST (Invitrogen) was used to replace the CMV promoter in plenti-CMV-gateway-3 \times HA-mPGK-GFP-P2A-Pur, plenti-CMV-3 \times HA-gateway-mPGK-Pur and plenti-CMV-gateway-3 \times HA-mPGK-Pur to generate plenti-Ubc-gateway-3 \times HA-mPGK-GFP-P2A-Pur, plenti-Ubc-3 \times HA-gateway-3 \times HA-mPGK-Pur, and plenti-Ubc-gateway-3 \times HA-mPGK-Pur. The EF1 α promoter from pEF-DEST51 (Invitrogen) was cloned into plenti-CMV-gateway-3 \times HA-mPGK-GFP-P2A-Pur, plenti-CMV-3 \times HA-gateway-mPGK-Pur and plenti-CMV-gateway-3 \times HA-mPGK-Pur to generate plenti-EF1 α -gateway-3 \times HA-mPGK-GFP-P2A-Pur, plenti-EF1 α -3 \times HA-gateway-3 \times HA-mPGK-Pur, and plenti-EF1 α -gateway-3 \times HA-mPGK-Pur.

Next, pDONR223-GSPT1 was shuttled into plenti-CMV-gateway-3 \times HA-mPGK-GFP-P2A-Pur and plenti-EF1 α -gateway-3 \times HA-mPGK-GFP-P2A-Pur via gateway LR (att L and att R) recombination. Similarly, pDONR223-GSPT1 wild-type and variants, pDONR223-SUP35, as well as pDONR223-GSPT1/SUP35 chimaeric fusion variants were shuttled into plenti-EF1 α -3 \times HA-mPGK-Pur. pDONR223-Flag-HA-CRBN and pDONR223-CRBN wild-type and mutants were cloned into plenti-CMV-gateway-3 \times HA-mPGK-GFP-P2A-Pur and plenti-EF1 α -gateway-3 \times HA-mPGK-P2A-Pur. pDONR223-IKZF1, pDONR223-IKZF1-Q146H and pDONR223-GSPT1 was cloned into plenti-Ubc-gateway-3 \times HA-mPGK-Pur or plenti-Ubc-3 \times HA-gateway-mPGK-Pur. A stuffer sequence was synthe-

sized *in vitro* by Invitrogen, and cloned into Mission TRC2 pLKO.5-pur empty vector control plasmid (Sigma) to generate pLKO.5pur-Stuffer. Coding sequence of 8 \times His-Ub was *in vitro* synthesized by Invitrogen and cloned into pcDNA3 to generate pcDNA3-8 \times His-Ub.

In-gel digestion and sample preparation for LC-MS/MS Analysis. Gel bands were excised and de-stained. Proteins were reduced and alkylated by addition of 50 mM DTT and 50 mM iodoacetamide before digestion by Trypsin at a final concentration of 25 ng μ L⁻¹ in 50 mM ammonium bicarbonate for 1 h on ice and additionally 16 h more at 37 °C using a shaking incubator to assure complete digestion. Digested tryptic peptides were extracted from gels by addition of 200 μ L of water and sonication for 10 min, washed in 5% formic acid in water and extracted in 50% acetonitrile in 5% formic acid in water, once in 70% acetonitrile and last in 100% acetonitrile. All extracted peptides were pooled together and were vacuum-dried and re-dissolved in 20 μ L of 0.1% TFA. Tryptic peptides were then concentrated and desalted using a Millipore C18 Zip Tip (Millipore). The eluted peptides were then vacuum-dried and re-dissolved in 100 μ L of LC-MS/MS loading buffer (2% acetonitrile in 0.1% formic acid in water).

One-dimensional LC-MS/MS analysis. Tryptic digested samples were subjected to an on-line analysis of peptides by high-resolution, high-accuracy LC-MS/MS, consisting of a Bruker-Michrom paradigm HPLC, a Zorbax C18 peptide trap column (Agilent technologies), a 15-cm Michrom Magic C18 column, a low-flow ADVANCED Michrom MS source, and a LTQ-Orbitrap XL (Thermo Fisher Scientific). The LC-MS/MS raw data were submitted to Sorcerer Enterprise v.3.5 release (Sage-N Research Inc.) with the SEQUEST algorithm to search the target-decoy ipi.Human.v3.73 database. The search results were viewed, sorted, filtered, and statically analysed by using comprehensive proteomics data analysis software, Peptide/Protein prophet v.4.02 (ISB). The differential spectral count analysis was done by QTools (Sanford Burnham).

Lentiviral shRNA. Pairs of synthetic complementary oligonucleotides targeting GSPT1 or a non-mammalian control (shCNTL) were annealed and cloned into pLKO.5pur-Stuffer. Targeting sequences are listed below: shCNTL, 5'-CAACAAGATGAAGAGCACCAA-3'; shGSPT1-1, 5'-GATTACCGTTTATTCCATA-3'; shGSPT1-2, 5'-CCGATGATGTAGAGACTGATA-3'; shGSPT1-3, 5'-CCTGCACAATACTGTGAGGAA-3'; shGSPT1-4, 5'-GCCGAACCTCAATA GATCAGT-3'.

Protein half-life analysis. 293FT HEK parental and CRBN^{-/-} cells were pretreated with DMSO or CC-885 (10 μ M) for 30 min, followed by the addition of 100 μ g mL⁻¹ cycloheximide (EMD) into the culture medium. At various time points, cells were collected and subjected to immunoblot analysis.

In vivo ubiquitination assay. The ubiquitination assays were carried out as described previously³⁴. Briefly, 293FT HEK parental or CRBN^{-/-} cells cultured in 10-cm plates were transfected with 12 μ g of pcDNA3-8 \times His-Ub. After 48 h, cells were treated with 10 μ M CC-885 and 10 μ M MG132. Eight hours later, cells were washed twice with ice cold PBS and resuspended in 1 ml PBS. 20 μ L of the cell suspension was boiled in LDS loading buffer (Thermo Fisher), and the remaining cells were collected via centrifugation and lysed in Buffer C (6 M guanidine-HCl, 0.1 M Na₂HPO₄/NaH₂PO₄, 20 mM imidazole (pH 8.0)). Next, whole-cell extracts were sonicated for 12 pulses, and mixed with 30 μ L of Ni-NTA agarose beads at room temperature for 4 h. Ni-NTA beads were then washed twice with Buffer C, twice with Buffer D (1 volume of Buffer C to 3 volumes of Buffer E), and twice with Buffer E (25 mM Tris HCl, 20 mM imidazole (pH 6.8)). Bound proteins were eluted by boiling in 2 \times LDS loading buffer and subjected to immunoblot analysis with two mouse anti-GSPT1 monoclonal antibodies 17G9 and 8A2 generated in house (Celgene), a mouse anti-His tag monoclonal antibody (Qiagen #34660, Qiagen) and a rabbit anti-human CRBN monoclonal antibody (CRBN65, Celgene). To detect the ubiquitination of GSPT1, boiled eluates were further diluted with 20 volumes of Buffer B (50 mM Tris (pH 7.4), 150 mM NaCl, 0.5% NP-40, one tablet of Complete ULTRA protease inhibitor cocktail (Roche)), followed by immunoprecipitation with a mouse anti-GSPT1 monoclonal antibody (17G9) immobilized on protein G beads. Beads were washed with Buffer B six times, and bound proteins were eluted by boiling in 1 \times SDS loading buffer and subjected to immunoblot analysis for GSPT1 with a rabbit anti-GSPT1 antibody (Abcam #49878) and two rabbit anti-ubiquitin antibodies (Cell Signaling #3933 and Bethyl #A300-317A).

In vitro ubiquitination assays. Purified E1, E2, ubiquitin, CUL4A-RBX1, cereblon-DDB1, and GSPT1 proteins were used to reconstitute the ubiquitination of GSPT1 *in vitro*. Human GSPT1 domains 2 and 3 (amino acids 437-633) was expressed in BL21 star DE3 *E. coli* and purified by maltose affinity resin (NEB), heparin HiTrap column affinity (GE Healthcare), and Superdex 75 16/60 size-exclusion chromatography (GE Healthcare), as described above. Human cereblon-DDB1 (cereblon amino acids 40-442 and full-length DDB1) and 6 \times his-DDB1 alone were co-expressed in SF9 insect cells in the presence of 50 μ M zinc acetate, and purified by nickel affinity resin (Qiagen), HiTrap ANX column ion exchange (GE Healthcare), and Sephacryl 400 16/60 size-exclusion chromatography

(GE healthcare), as described above. SV5-V and DDB1 were purified separately, incubated to form a complex, and the complex purified with size-exclusion chromatography as described above. Human full-length CUL4A and RBX1 were co-expressed in SF9 insect cells and purified by nickel affinity resin and Superdex 200 16/60 size-exclusion chromatography. Purified recombinant human Ube1 E1 (E-305), UbcH5a E2 (E2-616), and ubiquitin (U-100H) were purchased from R&D systems. Components were mixed to final concentrations of 1 μ M Ube1, 25 μ M UbcH5a, 200 μ M Ub, 1 μ M CUL4-RBX1, 25 μ M GSPT1, and 1 μ M cereblon-DDB1, 6 \times His-DDB1, or SV5-V-DDB1, with or without 100 μ M CC-885 in ubiquitination assay buffer (20 mM HEPES (pH 7.5), 150 mM NaCl, 10 mM MgCl_2). After preincubation for 10 min at room temperature, ubiquitination reactions were started by addition of ATP to a final concentration of 10 mM. Reactions were incubated at 30 °C for 2 h before separation by SDS-PAGE followed by coomassie staining or immunoblot analysis for GSPT1, ubiquitin, cereblon or SV5-V.

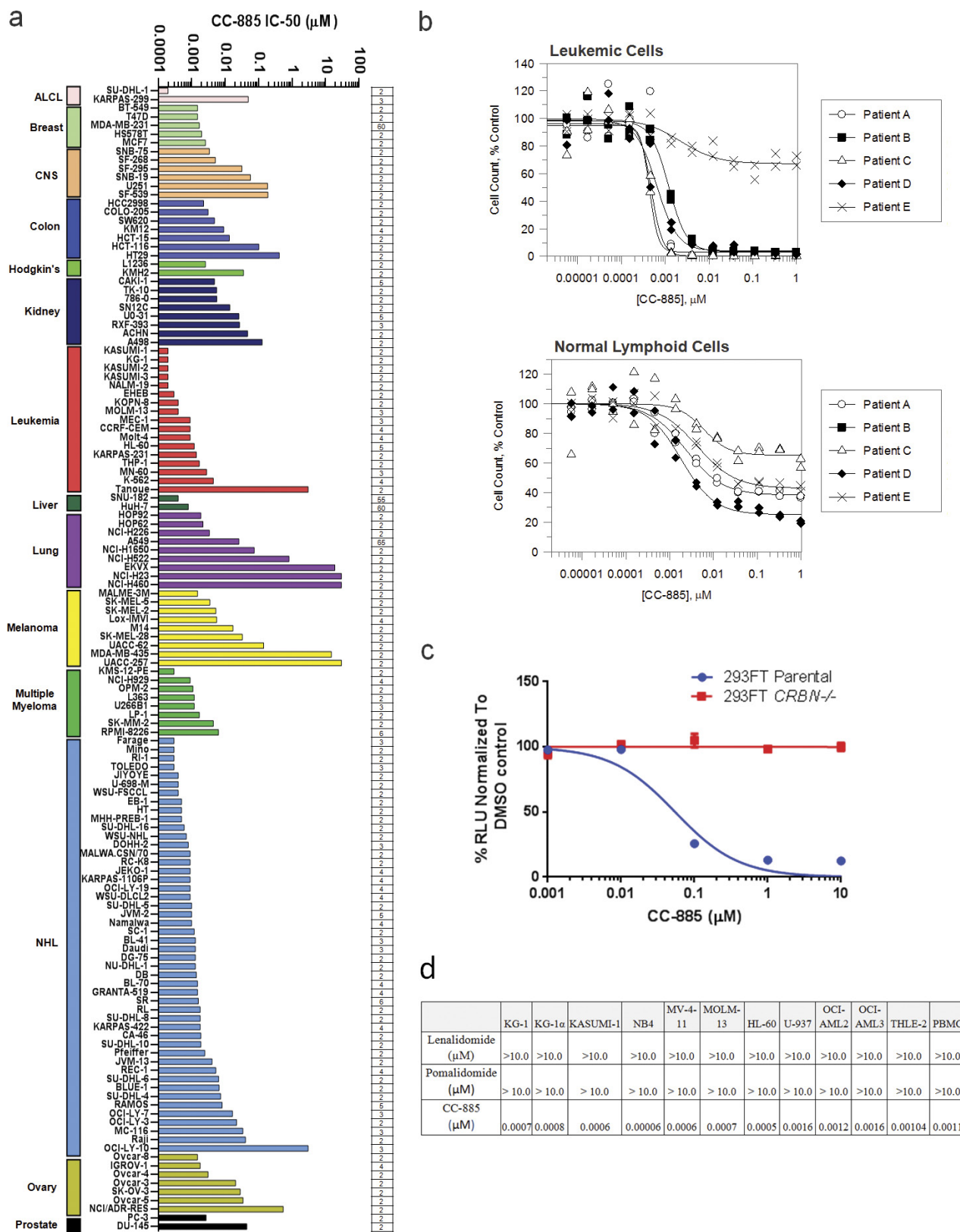
qRT-PCR Analysis. Following incubation with test compounds, cells were collected via centrifugation at 872g for 2 min. Cell pellets were then washed once in ice-cold PBS and snap-frozen in liquid nitrogen. Total RNA was extracted using RNeasy Mini Kit (Qiagen) according to manufacturer's instruction, and reverse-transcribed into first-strand cDNA using AffinityScript QPCR cDNA Synthesis Kit (Agilent) with random primer. Real-time PCR was performed in triplicate using ViiA 7 Real-Time PCR System with TaqMan Gene Expression Assay probes (Invitrogen) for GSPT1 (# 4331182) and GAPDH (#4326317E).

Lentiviral production and infection. Lentiviral plasmid was cotransfected with the 2nd Generation packaging system (ABM) into Lenti-X 293 cells using Lipofectamine 2000 (Thermo Fisher). After 12 h of incubation, media was changed

to fresh DMEM media supplemented with 20% FBS. At 36 h after transfection, viral supernatant was collected and fresh media was replenished. At 56 h, viral supernatant was collected again, combined with the first viral supernatant, cleared via centrifugation at 872g for 5 min, and then filtered through a 0.45 micron cellulose acetate or nylon filter unit. All lentiviral supernatants, except those of lentiviral shRNA vectors, were concentrated using a Lenti-X concentrator according to the manufacturer's instructions. AML cell lines were spin-inoculated with lentivirus at 872g for 90 min. After 12 h, viral supernatant was removed and complete culture media was added to the cells. 48 h later, cells were incubated with 1–2 μ g ml^{-1} puromycin or 10–20 μ g ml^{-1} blasticidin for an additional 3 days to select cells stably integrated with lentiviral vectors.

Data analysis. All cell proliferation assays were independently performed three times with three technical replicates for each sample. All qRT-PCR assays were carried out two times with three technical replicates for each sample. When possible, results were confirmed in multiple cell lines. For quantification of cell proliferation and qRT-PCR, data were analysed using GraphPad Prism and Excel, respectively. $P < 0.05$, unpaired two-sided t -test, is considered as significant.

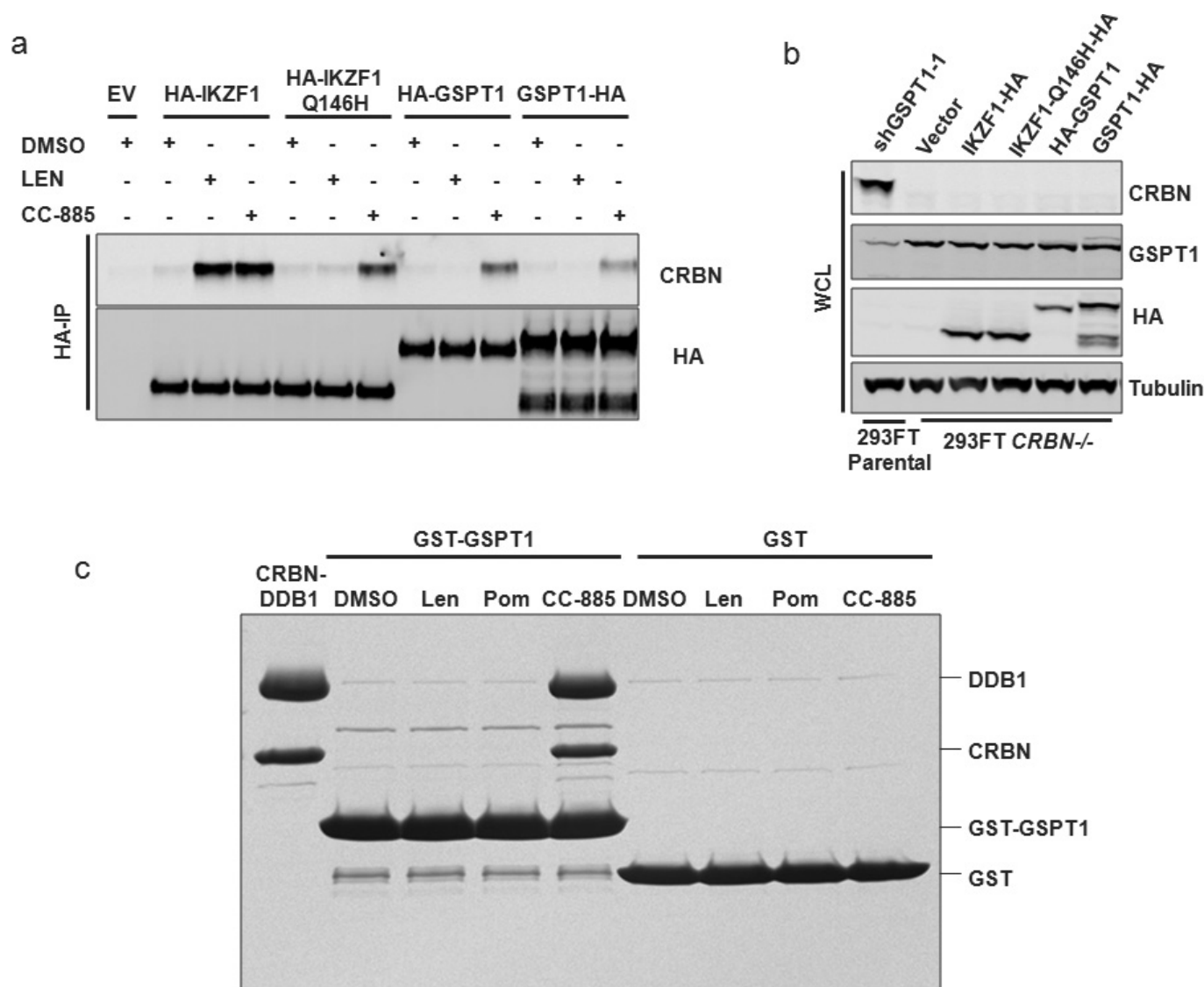
31. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
32. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
33. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67**, 355–367 (2011).
34. Bennett, T. A. *et al.* Pharmacological profiles of acute myeloid leukemia treatments in patient samples by automated flow cytometry: a bridge to individualized medicine. *Clin. Lymphoma Myeloma Leuk.* **14**, 305–318 (2014).



Extended Data Figure 1 | See next page for caption.

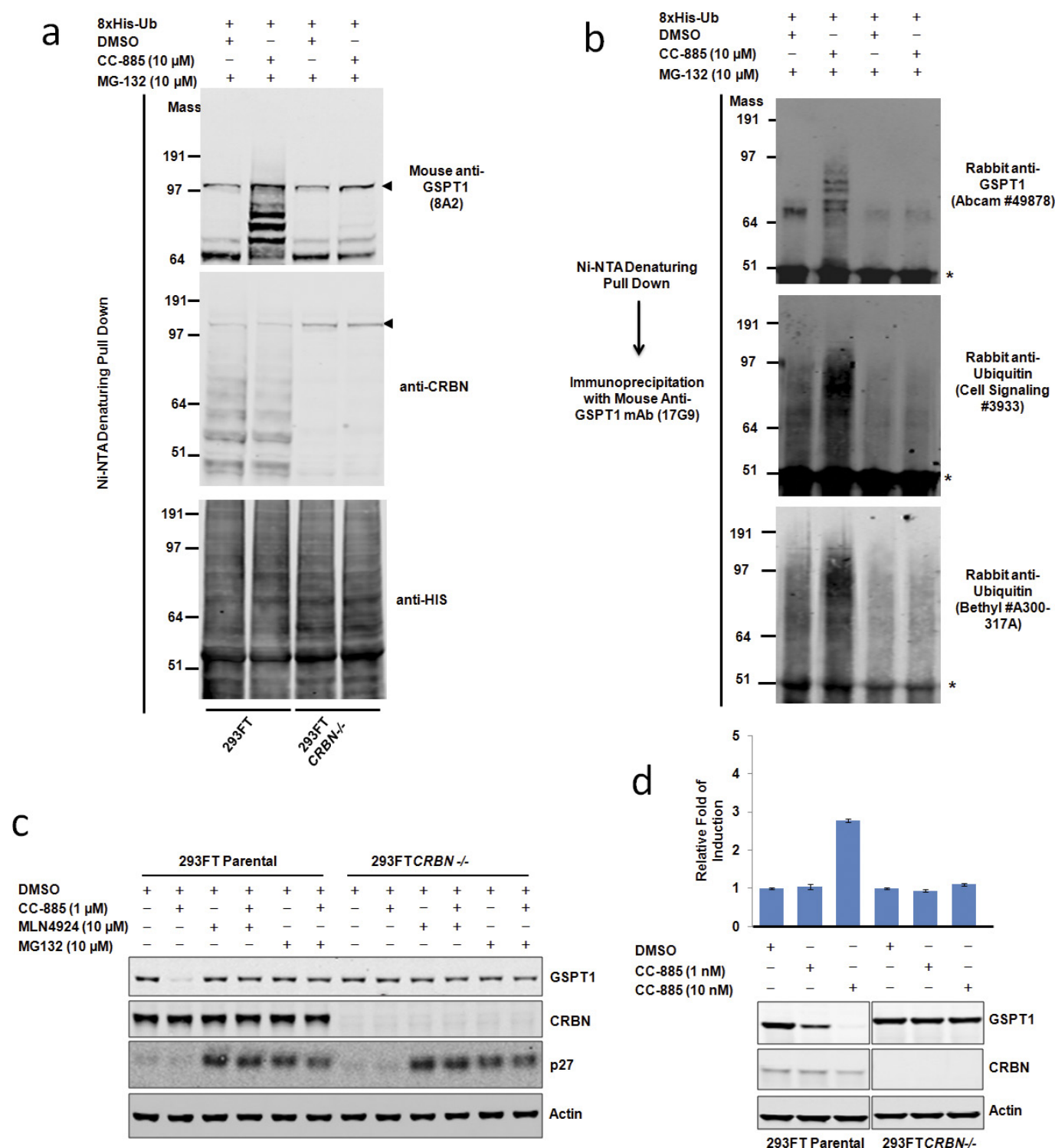
Extended Data Figure 1 | The anti-tumour effects of CC-885 are CRBN dependent. **a**, The growth inhibitory IC_{50} value of CC-885 in human cancer cell lines. The cell proliferation of 132 human cancer cell lines treated with varying concentrations of CC-885 for 3 days was assessed by Cell Titre Glo assay. Cancer types as indicated at the bottom are labelled with different colours. The IC_{50} value of growth inhibition was determined using ActivityBase (IDBS). Data are the mean of two or more biological replicates. The n number for each cell line is shown on the right of the panel. **b**, The effect of CC-885 on AML samples taken from patients. The top panel shows the effects on leukaemic cells; the bottom panel shows the effects on normal lymphoid cells. The sample from patient B did not

contain sufficient normal lymphoid cells for analysis. **c**, The effect of CC-885 on cell proliferation in parental and *CRBN*^{-/-} 293FT HEK cells. Result is representative of three biological replicates. Data are shown as mean \pm s.d., $n = 3$ technical replicates. **d**, The effect of lenalidomide, pomalidomide and CC-885 on cell proliferation in AML cell lines, THLE-2 and human PBMC. Cells were treated with varying concentrations of lenalidomide, pomalidomide or CC-885. At day 3, cell proliferation was assessed using CTG assay. Numbers shown are the growth inhibitory IC_{50} values of lenalidomide, pomalidomide and CC-885, from biological replicates with $n = 5, 3$, and 20, respectively.



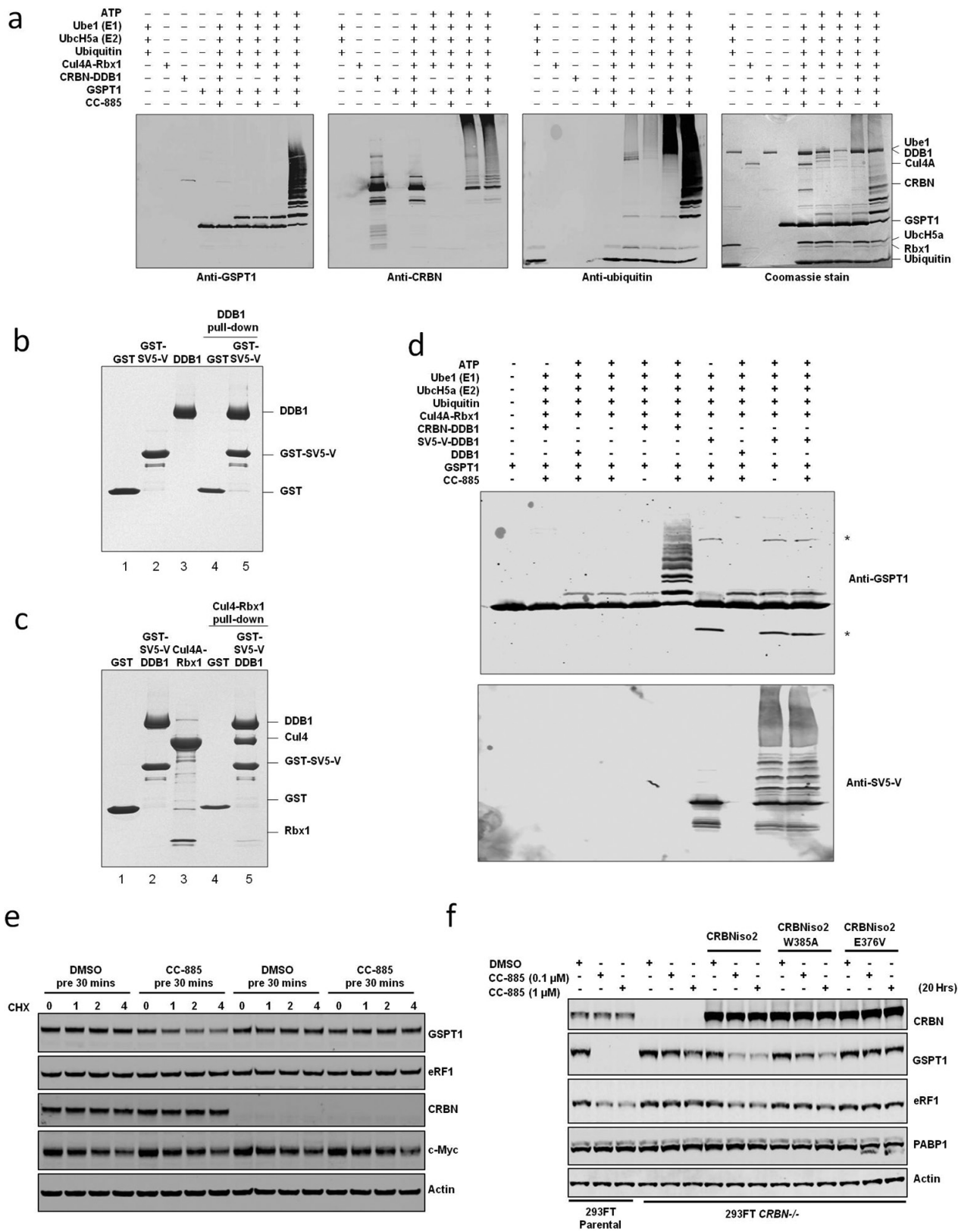
Extended Data Figure 2 | CC-885, but not lenalidomide, promotes the interaction between GSPT1 and CRBN *in vitro*. **a, b,** Immunoblot analysis of anti-HA immunoprecipitates (**a**) and whole-cell lysates (WCL) (**b**). HA-tagged GSPT1 or IKZF1 produced in 293FT CRBN^{-/-} HEK cells were used to capture CRBN from 293FT HEK cells expressing GSPT1-specific shRNA, shGSPT1-1. DMSO, 10 μ M lenalidomide or 10 μ M CC-885 were included in the binding assay. The IKZF1 Q146H mutation results in a reduction of cereblon binding mediated by either lenalidomide or CC-885. Some residual binding is observed with CC-885, consistent with Fig. 1d, which shows that CC-885 is more potent than lenalidomide

against IKZF1. **c,** Coomassie stain of CRBN-DDB1 pull-down with GSPT1 using purified components. Purified GST or GST-GSPT1 domains 2 and 3 (amino acids 437–633) bound to magnetic glutathione beads was incubated with purified CRBN-DDB1 in the presences of CC-885, lenalidomide (len), pomalidomide (pom), or DMSO (vehicle) for 1 h at room temperature before three rapid washes. CC-885, but not lenalidomide, pomalidomide or DMSO, mediated the binding of GST-GSPT1 to CRBN-DDB1. This experiment was performed three times. For gel source data, see Supplementary Information Fig. 1.



Extended Data Figure 3 | Transcriptional and post-transcriptional regulation of GSPT1 by CC-885. **a**, Ubiquitinated protein products, as shown in Fig. 1e, were blotted for GSPT1 with mouse anti-GSPT1 monoclonal antibody 8A2, CRBN, and His-tagged Ubiquitin. Arrowheads mark nonspecific bands. **b**, 293FT parental and *CRBN*^{-/-} cells transiently expressing His-tagged ubiquitin was treated with CC-885 and MG132 as indicated for 8 h. Ubiquitinated protein products were pulled down with Ni-NTA agarose under denaturing conditions, followed by elution via boiling in SDS loading buffer. Eluates were diluted with immunoprecipitation buffer and GSPT1 was immunoprecipitated with

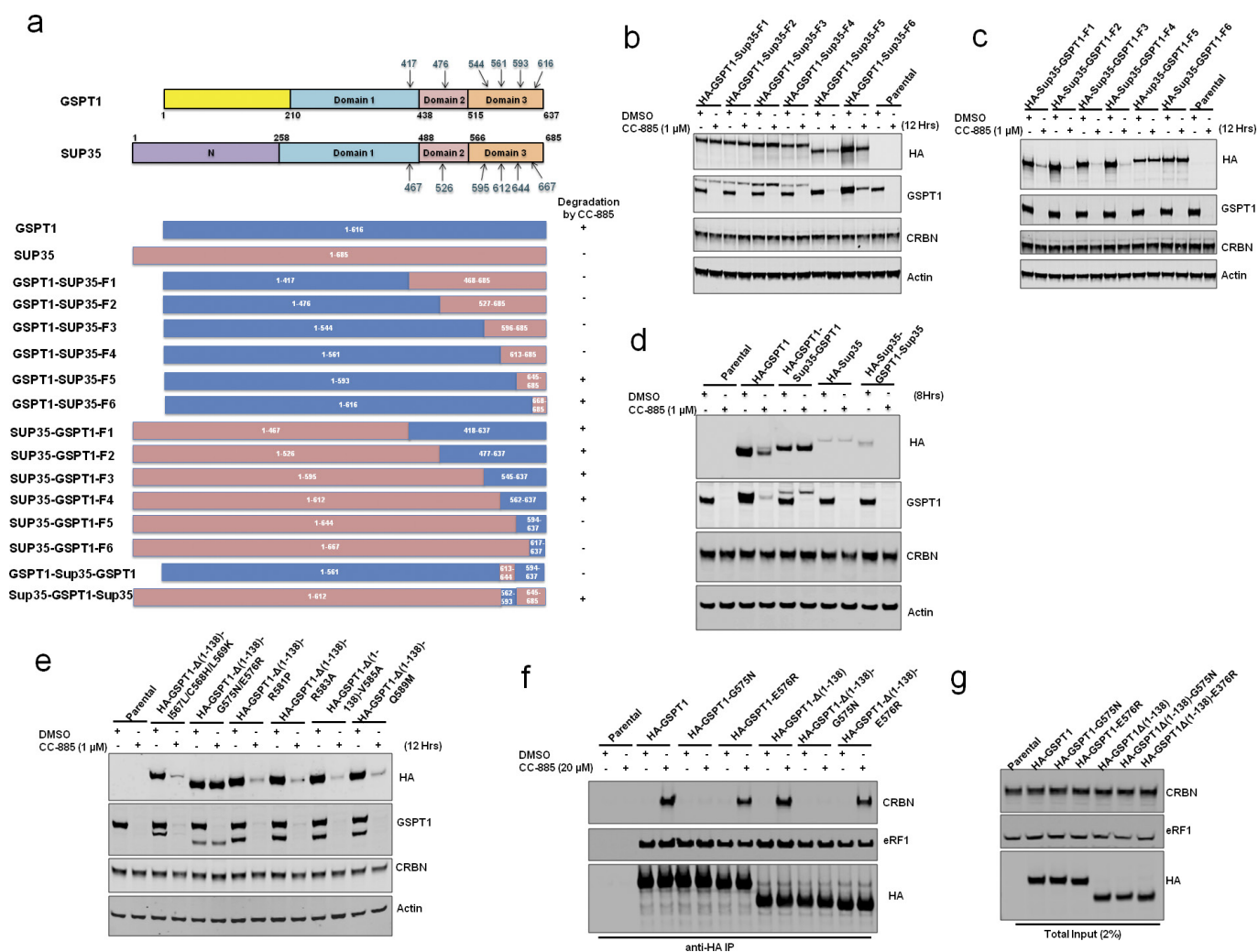
a mouse anti-GSPT1 monoclonal antibody 17G9. Immunoprecipitates were then subjected to immunoblot analysis for GSPT1 and ubiquitin with two different rabbit anti-ubiquitin antibodies. **c**, Immunoblot analysis of 293FT parental and *CRBN*^{-/-} cells treated with CC-885 for 4 h. Cells were pretreated with DMSO, CC-885, MLN4924 or MG132 as indicated. **d**, 293FT parental or *CRBN*^{-/-} cells treated with DMSO or CC-885 for 24 h were subjected to real-time qPCR (top) and immunoblot analysis (bottom). Result is representative of two biological replicates. Data are mean \pm s.d., $n = 3$ technical replicates. For gel source data, see Supplementary Information Fig. 1.



Extended Data Figure 4 | See next page for caption.

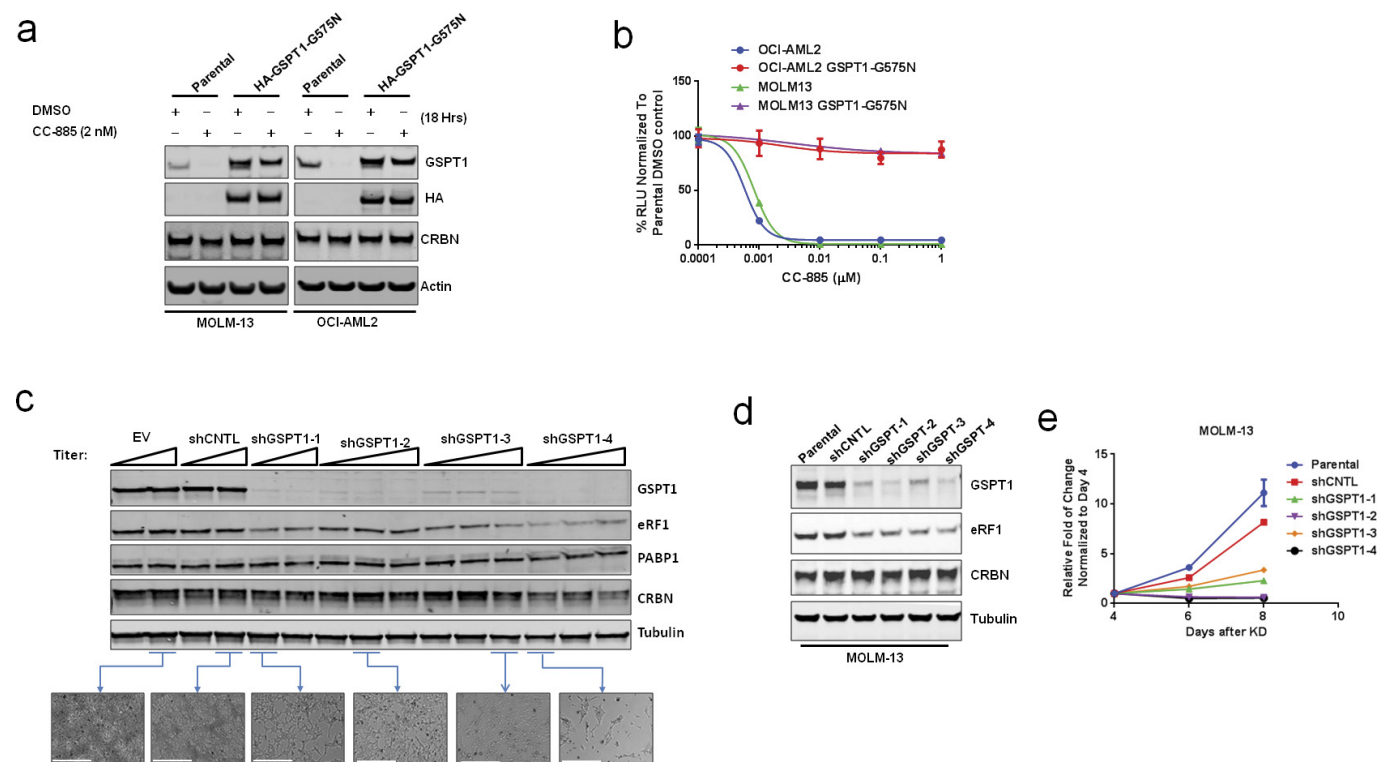
Extended Data Figure 4 | The CC-885 induced GSPT1 ubiquitination and degradation relies on CRBN. **a**, Controls for the *in vitro* ubiquitination assay shown in Fig. 1f. Recombinant protein products as indicated were incubated with or without ATP (10 mM) and CC-885 (100 μ M) in the ubiquitination assay buffer at 30 °C for 2 h, and then separated by SDS–PAGE followed by Coomassie staining or immunoblot analysis with anti-GSPT1, anti-cereblon, and anti-ubiquitin antibodies. The anti-GSPT1 blot shown is the same one shown in Fig. 1f. The additional blots and Coomassie staining analysis was performed on samples from the same *in vitro* ubiquitination reactions. Anti-GSPT1 blotting and Coomassie staining was performed three times, anti-cereblon and anti-ubiquitin blotting was performed twice. **b–d**, To demonstrate the CRBN-dependence of GSPT1 ubiquitination by CC-885–CRL4, we reconstituted DDB1 with an alternative DCAF, SV5-V, and showed that SV5-V is not capable of recruiting GSPT1 for polyubiquitination. **b**, DDB1 binds to GST–SV5-V, but not GST alone. Coomassie stain with lanes 1, 2, and 3 showing individual proteins, and lanes 4 and 5 showing pull-down of DDB1 incubated with GST or GST–SV5-V bound to glutathione magnetic beads and washed three times. This experiment was performed twice. **c**, GST–SV5-V–DDB1 protein complex generated by pull-down in **b** was used to show binding of CUL4–RBX1 to the GST–SV5-V–DDB1 protein complex but not GST alone. Coomassie stain with lanes 1, 2, and 3 showing individual proteins or protein complexes used in the pull-down, and lanes 4 and 5 showing pull-down of CUL4–RBX1 incubated with GST or GST–SV5-V–DDB1 bound to glutathione magnetic beads and washed three times. This experiment was performed once. **d**, *In vitro* ubiquitination of GSPT1 by CRBN–DDB1 but not SV5-V–DDB1. SV5-V–DDB1 complex was formed by incubation

of individually purified proteins and purification over size-exclusion chromatography. Recombinant protein products as indicated were incubated with either CRBN–DDB1, SV5-V–DDB1, DDB1 alone, or the absence of any DDB1 complex, with and without ATP (10 mM) or CC-885 (100 μ M) in ubiquitination assay buffer at 30 °C for 2 h, and then separated by SDS–PAGE followed by immunoblot analysis with anti-GSPT1 and anti-SV5-V antibodies. Asterisk indicates background bands present from the SV5-V protein purification. The anti-SV5-V western blot indicates that SV5-V is auto-ubiquitinated, as expected for a functional DCAF that is bound to the DDB1–CUL4–RBX1 complex. **e**, Immunoblot analysis of 293FT parental and *CRBN*^{−/−} cells treated with 100 μ g ml^{−1} cyclohexamide (CHX) for the indicated periods. Where indicated, cells were pretreated with DMSO or 10 μ M CC-885 for 30 min. This experiment was performed once. **f**, Immunoblot analysis of 293FT parental cells, *CRBN*^{−/−} cells, and *CRBN*^{−/−} cells stably expressing CRBN wild-type or variants as indicated. CRBNiso2 (CRBN isoform 2) showed similar E3 ligase activity towards GSPT1 as compared to CRBN isoform 1 (data not shown). CRBN isoform 2 lacks an alanine residue at position 23 of CRBN isoform 1, and as such the numbering is shifted compared to isoform 1. Note that the CRBN(W385A) mutant showed diminished activity in cells treated with 0.1 μ M CC-885, and similar activity in cells treated with 1 μ M CC-885, compared to wild-type CRBN. CRBN(E376V) mutant had no activity at both concentrations. Immunoblot analysis of 293FT parental and *CRBN*^{−/−} cells treated with 100 μ g ml^{−1} cyclohexamide for the indicated periods. Where indicated, cells were pretreated with DMSO or 10 μ M CC-885 for 30 min. For gel source data, see Supplementary Information Fig. 1.



Extended Data Figure 5 | Identification of the region of GSPT1 indispensable for CC-885-dependent destruction. **a**, Schematic of human GSPT1 and *Saccharomyces cerevisiae* SUP35 chimaeras. Truncation analysis of GSPT1 revealed that domains 2 and 3 of GSPT1 contain the CC-885-dependent CRBN-binding motif (data not shown). In this region, GSPT1 and SUP35 share 53% sequence identity. Chimaeric fusion sites were selected from several stretches of identical regions to ensure proper folding of the resultant fusion product. Changes in protein level of these fusion proteins in 293T HEK cells in response to CC-885 treatment, as shown in **b–d**, is summarized on the right. +, protein degraded; –, no change of protein level. **b–d**, Immunoblot analysis of 293T HEK parental

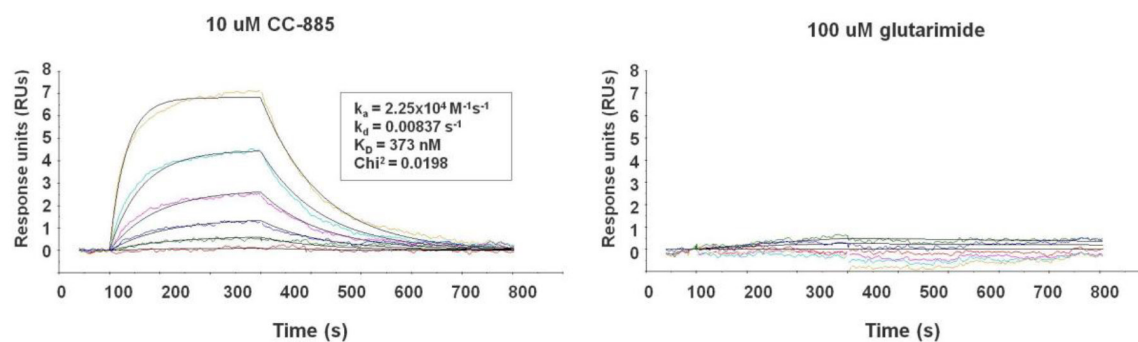
cells or cells stably expressing HA-tagged GSPT1, SUP35 or GSPT1–SUP35 chimaeric proteins. Note that SUP35 expression level is relatively low compared to GSPT1, possibly owing to its intrinsic instability when expressed in human cell lines. **e**, Immunoblot analysis of 293T HEK cells stably producing HA-tagged GSPT1 wild-type and variants. Non-conserved amino acids in GSPT1 as shown in Fig. 2a were replaced with corresponding residues in SUP35. **f, g**, Immunoblot analysis of anti-HA immunoprecipitates (**f**) and whole-cell extracts (**g**) of 293T HEK cells expressing GSPT1 wild-type and mutants. DMSO or CC-885 was added into the whole-cell extract after lysis for 6 h. For gel source data, see Supplementary Information Fig. 1.



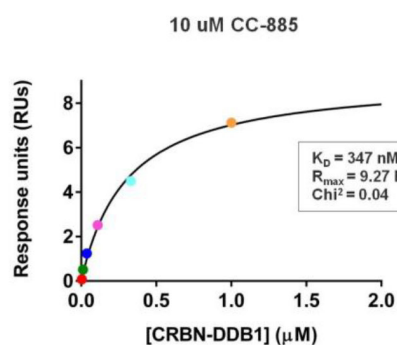
Extended Data Figure 6 | Loss of GSPT1 is the cause of growth inhibition induced by CC-885. **a**, Immunoblot analysis of MOLM-13 and OCI-AML2 parental cells or cells stably expressing HA-tagged GSPT1(G575N). Cells were treated with CC-885. **b**, Cells shown in **a** were incubated with CC-885 at the indicated concentration for 3 days. Cell proliferation was determined by CellTiter-Glo (CTG). Result is representative of three biological replicates. Data are presented as mean \pm s.d., $n = 3$ technical replicates. **c**, 293T HEK cells were infected with an increased amount of empty lentiviral vector (EV) or vectors expressing control shRNA or any of the four GSPT1-specific shRNAs. Seven days after infection, cells were imaged using phase-contrast

microscope (bottom) and collected for immunoblot analysis (top). Scale bars measure 0.5 mm. Images shown are representative of three captured images. **d**, **e**, Cell growth curve of MOLM-13 parental cells or cells transduced with lentiviral vectors expressing a control shRNA (shCNTL) or GSPT1 specific shRNA (shGSPT1-1 to 4). Immunoblot analyses of whole-cell extracts (**d**) were carried out at day 4 after transduction with lentiviral vectors. Cell growth (**e**) was quantified with CTG at day 4, 6, and 8 after transduction. Result is representative of three biological replicates. Data are mean \pm s.d., $n = 3$ technical replicates. For gel source data, see Supplementary Information Fig. 1.

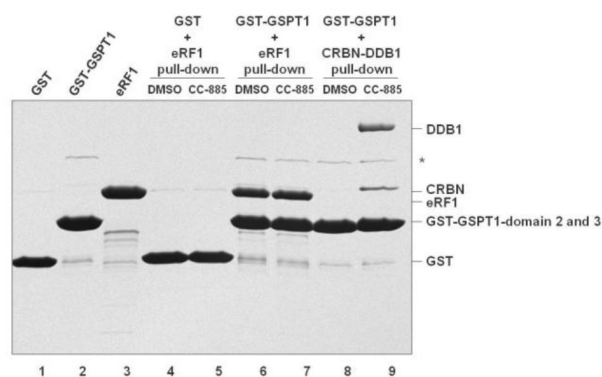
a



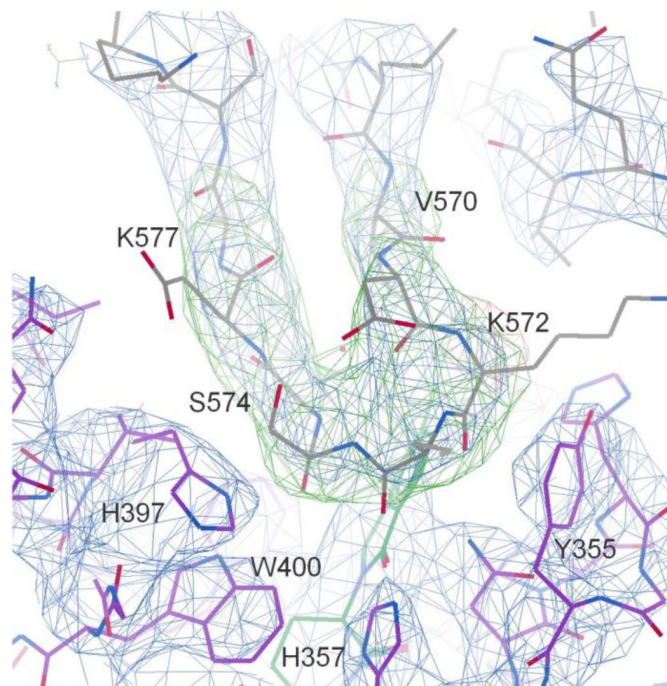
b



c



d

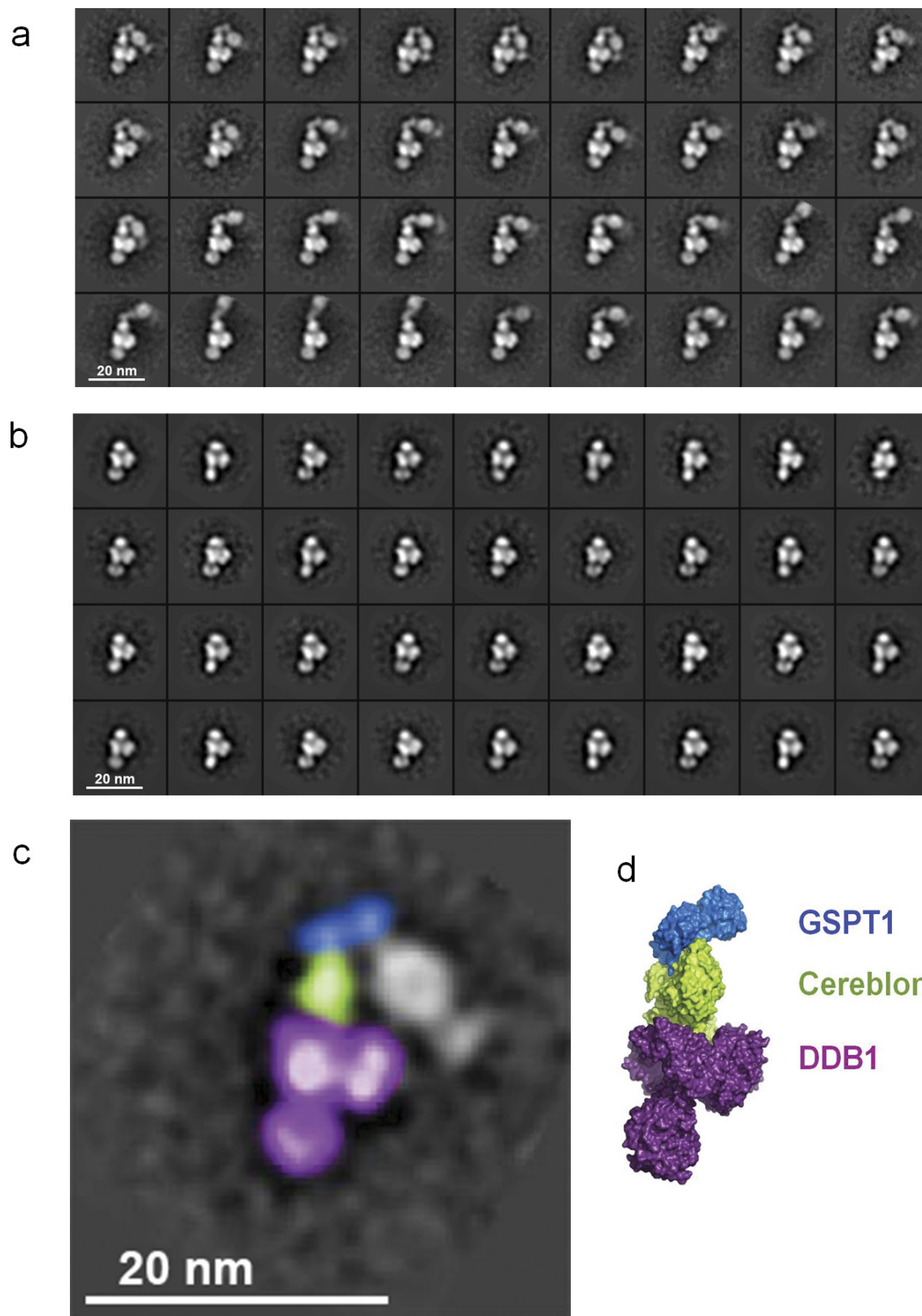


Extended Data Figure 7 | See next page for caption.

Extended Data Figure 7 | Kinetic parameters for cereblon–DDB1–CC-885 binding to GSPT1, and sample electron density from the crystal structure of the complex.

a, Reference-corrected surface plasmon resonance binding curves for various concentrations of cereblon–DDB1 (threefold dilutions from 1 μ M, coloured traces) flowed over a surface of covalently immobilized anti-GST antibody bound to GST–GSPT1 domains 2 and 3 at 10 °C in the presence of saturating levels of CC-885 or control compound glutarimide. Kinetic parameters shown were determined by fitting with a 1:1 kinetic binding model (black lines) using the Biacore T200 kinetic analysis software package. Binding in the presence of glutarimide could not be quantified. We show a representative set of curves from three independent experiments. For a 1:1 binding model where $A + B \rightleftharpoons AB$, the net rate of complex formation is given by the equation $d[AB]/dt = k_a[A][B] - k_d[AB]$ and the rate of complex disassociation is given by $k_d[AB]$, where k_a is the association rate constant ($M^{-1}s^{-1}$) and k_d is the dissociation rate constant (s^{-1}). K_D , the equilibrium dissociation constant, is defined by $K_D = k_d/k_a$. R_{max} is a measurement of the analyte binding capacity, or maximum response. **b**, Analysis of the steady-state response versus the concentration of analyte in the presence

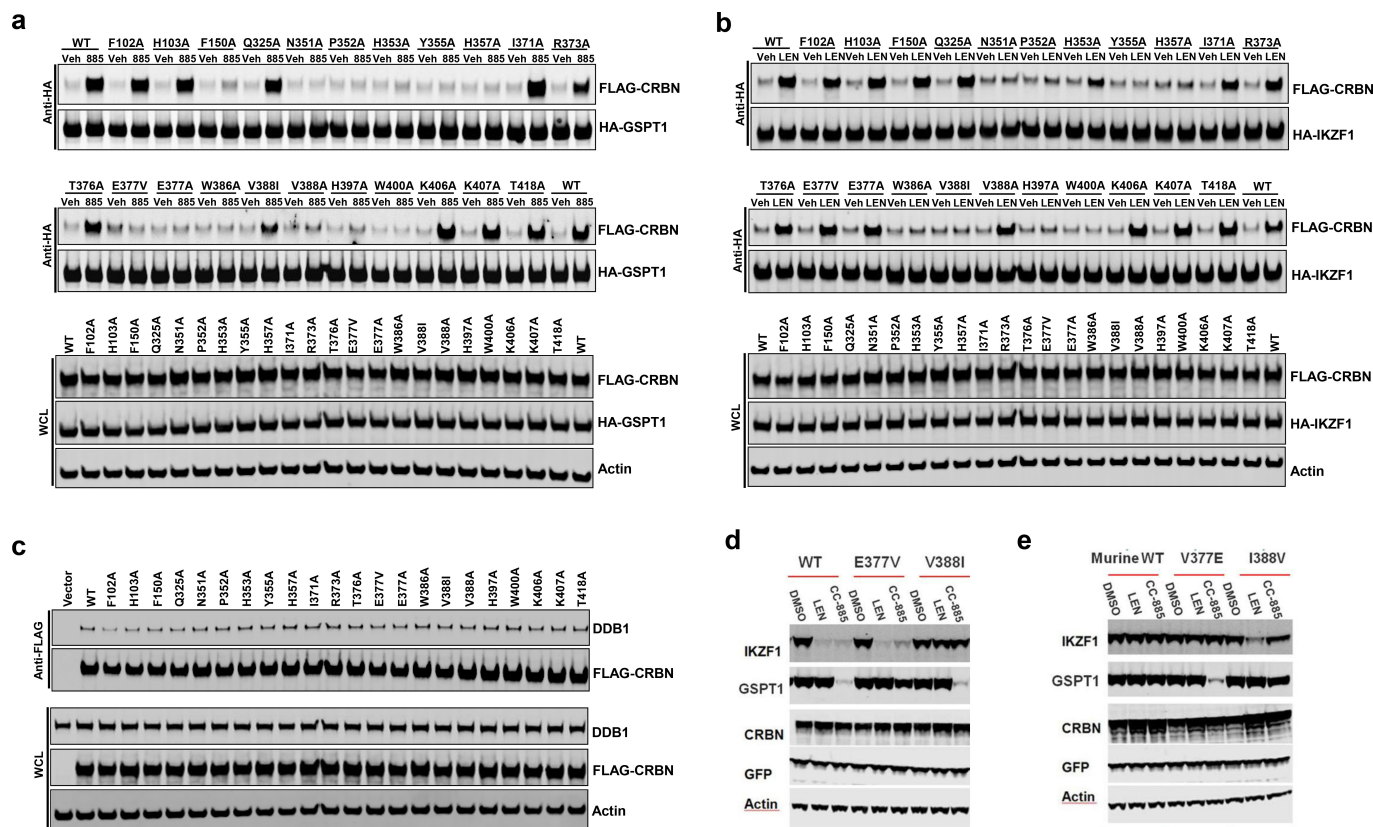
of CC-885, and the determined affinity constant. **c**, GST–GSPT1 interacts with endogenous binding partner eRF1. We show that the purified GST–GSPT1 domains 2 and 3 (amino acids 437–633) protein used in this SPR binding assay and the electron microscopy experiments (Extended Data Fig. 8) is competent to bind purified eRF1 as reported in ref. 17. Coomassie stain, with lanes 1, 2 and 3 showing individual proteins. Lanes 4 and 5 show pull-down of GST; lanes 6 and 7 show pull-down of GST–GSPT1 domains 2 and 3, all were bound to magnetic glutathione beads incubated with eRF1 \pm CC-885 for 1 h and washed three times. Lanes 8 and 9 show pull-down of GST–GSPT1 bound to magnetic glutathione beads incubated with CRBN–DDB1 \pm CC-885 for 1 h and washed three times. This experiment was performed twice. For gel source data, see Supplementary Information Fig. 1. **d**, Sample electron density from the cereblon–DDB1–CC-885–GSPT1 crystal structure with cereblon residues shown in purple, GSPT1 residues shown in grey, and CC-885 shown in green. Refined $2F_o - F_c$ density is shown as a blue mesh contoured at 1.4σ . $F_o - F_c$ difference density, shown as a green mesh contoured at 4σ , was generated by a single round of Refmac5 refinement calculated in the absence of GSPT1 residues 570–577.



Extended Data Figure 8 | See next page for caption.

Extended Data Figure 8 | Imaging of GSPT1 binding to cereblon–DDB1–CC-885 by negative stain electron microscopy. **a**, Negative stain class averages of N-terminal-tagged GST–GSPT1 domain 2 and 3 (amino acids 437–633) bound to purified cereblon–DDB1 in the presence of CC-885. As the GST-tag was fused to the N terminus of domain 2 of GSPT1, domain 3 can be identified as mediating the interaction with cereblon. Whereas the GST tag appears flexible in position, the two domains of GSPT1 are consistent in their orientation with the cereblon–DDB1 complex. GST-dimerization mediates the binding of a second substrate in the majority of the classes. Each class average is composed of between 40 and 70 individual particles. This experiment was performed three times. **b**, Negative stain class averages of cereblon–DDB1 and GST–GSPT1 domains 2 and 3 (amino acids 437–633) in the presence of DMSO instead

of CC-885. No classes containing bound substrate were observed in the absence of CC-885. Each class average is composed of between 40 and 70 individual particles. This experiment was performed once. **c**, A negative stain class average of GST–GSPT1 domains 2 and 3 bound to cereblon–DDB1–CC-885, with DDB1 shaded purple, cereblon shaded green, GSPT1 domains 2 and 3 shaded blue, and the second dimerized GST tag and second GSPT1 left uncoloured. **d**, For comparison, the crystal structure of GSPT1 domains 2 and 3 bound to cereblon–DDB1–CC-885 with DDB1 in purple, cereblon in green, and GSPT1 in blue. The electron micrographs revealed a consistent configuration of GSPT1 with cereblon in all complex class averages and confirmed that domain 3 mediates cereblon interactions on the basis of the orientation of the GST-tag.



Extended Data Figure 9 | Effects of cereblon surface mutations on substrate binding. Cereblon and substrate proteins were co-expressed in 293FT *CRBN*^{-/-} cells, co-immunoprecipitated in the presence or absence of CC-885 and lenalidomide, and analysed by western blot. **a**, Co-immunoprecipitation of GSPT1 with wild-type and mutant cereblon. veh, vehicle, DMSO; 885, 10 μ M CC-885. **b**, Co-immunoprecipitation of Ikaros with wild-type and mutant cereblon. veh, vehicle, DMSO; LEN, 10 μ M lenalidomide. **c**, Co-immunoprecipitation of DDB1 with

wild-type and mutant cereblon. Results are representative of three biological replicates. **d**, **e**, Western blots showing the effect of lenalidomide or CC-885 on Ikaros and GSPT1 degradation; **d**, effect with human cereblon, **e**, effect with mouse cereblon. For convenience, the human amino acid numbering is used to discuss the corresponding residues in mouse. GFP and actin are shown as transfection and loading controls, respectively. This is a representative experiment of three biological replicates. For gel source data, see Supplementary Information Fig. 1.

Extended Data Table 1 | Data collection and refinement statistics from the cereblon–DDB1–CC-885–GSPT1 crystal structure

CRBN-DDB1-CC-885-GSPT1	
Data collection	
Space group	<i>P</i> 2
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	156.8 111.5 175.1
α , β , γ (°)	90 95.8 90
Resolution (Å)	50-3.6 (3.66-3.6)*
<i>R</i> _{merge} (%)	19.8 (79.6)
<i>I</i> / σ <i>I</i>	10.6 (2.1)
Completeness (%)	94.4 (93.9)
Redundancy	4.8 (4.5)
Refinement	
Resolution (Å)	50.0-3.6
No. reflections	62827
<i>R</i> _{work} / <i>R</i> _{free}	0.224 / 0.273
No. atoms	
Protein	25202
Ligand/ion	62/2
Water	0
<i>B</i> factors	
Protein	117.9
Ligand/ion	99.9/167.6
Water	N/A
r.m.s. deviations	
Bond lengths (Å)	0.017
Bond angles (°)	1.796

*Values in parentheses are for highest-resolution shell.

Imaging the water snow–line during a protostellar outburst

Lucas A. Cieza^{1,2}, Simon Casassus^{2,3}, John Tobin⁴, Steven P. Bos⁴, Jonathan P. Williams⁵, Sebastian Perez^{2,3}, Zhaohuan Zhu⁶, Claudio Caceres^{2,7}, Hector Canovas^{2,7}, Michael M. Dunham⁸, Antonio Hales⁹, Jose L. Prieto^{1,10}, David A. Principe^{1,2}, Matthias R. Schreiber^{2,7}, Dary Ruiz-Rodriguez¹¹ & Alice Zurlo^{1,2,3}

A snow-line is the region of a protoplanetary disk at which a major volatile, such as water or carbon monoxide, reaches its condensation temperature. Snow-lines play a crucial role in disk evolution by promoting the rapid growth of ice-covered grains^{1–6}. Signatures of the carbon monoxide snow-line (at temperatures of around 20 kelvin) have recently been imaged in the disks surrounding the pre-main-sequence stars TW Hydra^{7–9} and HD163296 (refs 3, 10), at distances of about 30 astronomical units (AU) from the star. But the water snow-line of a protoplanetary disk (at temperatures of more than 100 kelvin) has not hitherto been seen, as it generally lies very close to the star (less than 5 AU away for solar-type stars¹¹). Water-ice is important because it regulates the efficiency of dust and planetesimal coagulation⁵, and the formation of comets, ice giants and the cores of gas giants¹². Here we report images at 0.03-arcsec resolution (12 AU) of the protoplanetary disk around V883 Ori, a protostar of 1.3 solar masses that is undergoing an outburst in luminosity arising from a temporary increase in the accretion rate¹³. We find an intensity break corresponding to an abrupt change in the optical depth at about 42 AU, where the elevated disk temperature approaches the condensation point of water, from which we conclude that the outburst has moved the water snow-line. The spectral behaviour across the snow-line confirms recent model predictions¹⁴: dust fragmentation and the inhibition of grain growth at higher temperatures results in soaring grain number densities and optical depths. As most planetary systems are expected to experience outbursts caused by accretion during their formation^{15,16}, our results imply that highly dynamical water snow-lines must be considered when developing models of disk evolution and planet formation.

V883 Ori is an FU Orionis (FU Ori) type star that was identified as such¹⁷ via follow-up spectroscopy of deeply embedded sources from the Infrared Astronomical Satellite. It is located in the Orion Nebula Cluster, which is at a distance of 414 ± 7 parsecs from Earth¹⁸. The mass of V883 Ori's protoplanetary disk is greater than about $0.3M_{\odot}$ (where M_{\odot} is the mass of the Sun), and its bolometric luminosity is $400L_{\odot}$ (ref. 19). We have obtained 230 GHz/1.3 mm (band-6) observations of V883 Ori using the Atacama Large Millimeter/Submillimeter Array (ALMA), in four different array configurations with baselines ranging from 14 metres to 12.6 kilometres, taken in ALMA cycles 2 and 3. These new ALMA observations include continuum and the ^{12}CO , ^{13}CO , and $\text{C}^{18}\text{O } J=2-1$ spectral lines. We use the C^{18}O gas line to investigate the dynamics of the system at $0.2''$ (90 AU) resolution, and the continuum data to constrain the physical properties of the dust in the V883 Ori disk at $0.03''$ (12 AU) resolution. In Fig. 1a we show our cycle-3 continuum image at $0.03''$ resolution—the highest resolution ever obtained

for a FU Ori object at millimetre wavelengths. We find that the V883 Ori disk has a two-region morphology, with a very bright inner disk (radius $\sim 0.1''$, 42 AU) and a much more tenuous outer disk extending

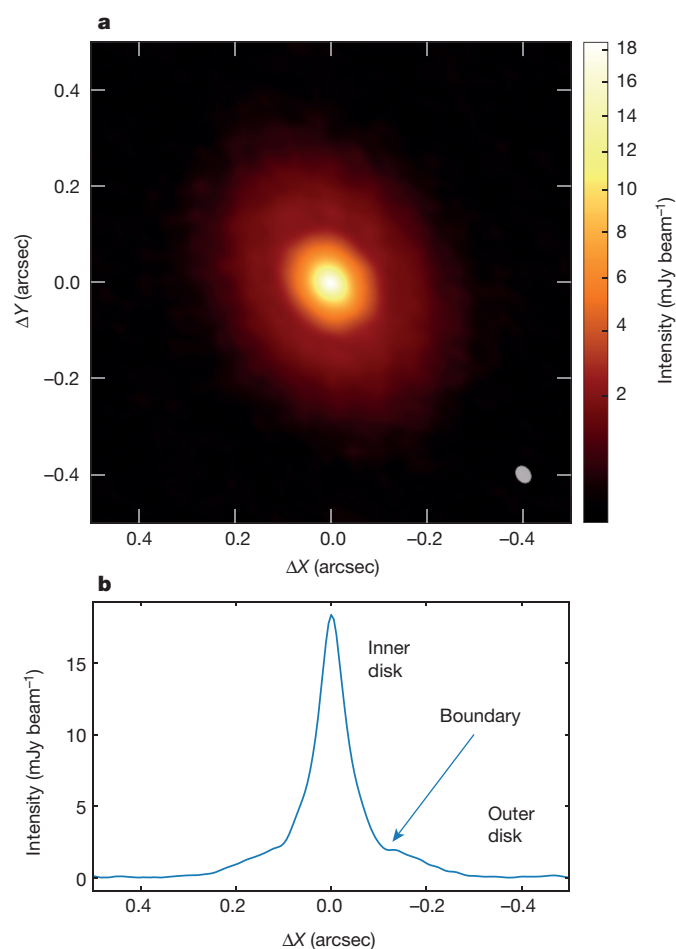


Figure 1 | ALMA observations of V883 Ori. **a**, The band-6 image at $0.03''$ (12 AU) resolution obtained on 27 October 2015. **b**, The intensity profile along the major axis. There is a very bright inner disk with radius $\sim 0.1''$ (42 AU), surrounded by a much more tenuous outer disk extending out to radius $\sim 0.3''$ (125 AU). The boundary between these two regions is sharp and probably unresolved. X and Y are the right ascension and the declination, respectively.

¹Núcleo de Astronomía, Facultad de Ingeniería, Universidad Diego Portales, Av. Ejercito 441, Santiago 8370191, Chile. ²Millennium Nucleus 'Protoplanetary Disks in ALMA Early Science', Av. Ejercito 441, Santiago 8370191, Chile. ³Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago 8330015, Chile. ⁴Leiden Observatory, Leiden University, PO Box 9513, 2300RA Leiden, The Netherlands. ⁵Institute for Astronomy, University of Hawaii at Manoa, Woodlawn Drive, Honolulu, Hawaii 96822, USA. ⁶Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Peyton Hall, Princeton, New Jersey 08544, USA. ⁷Departamento de Física y Astronomía, Universidad Valparaíso, Av. Gran Bretaña 111, Valparaíso 2373195, Chile. ⁸Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA. ⁹Joint ALMA Observatory, Alonso de Córdova 3107, Vitacura, Santiago 7630355, Chile. ¹⁰Millennium Institute of Astrophysics, Av. Vicuña Mackenna 4860, Macul, Santiago 7820436, Chile. ¹¹Research School of Astronomy and Astrophysics, Australian National University, Canberra, Australian Capital Territory 2611, Australia.

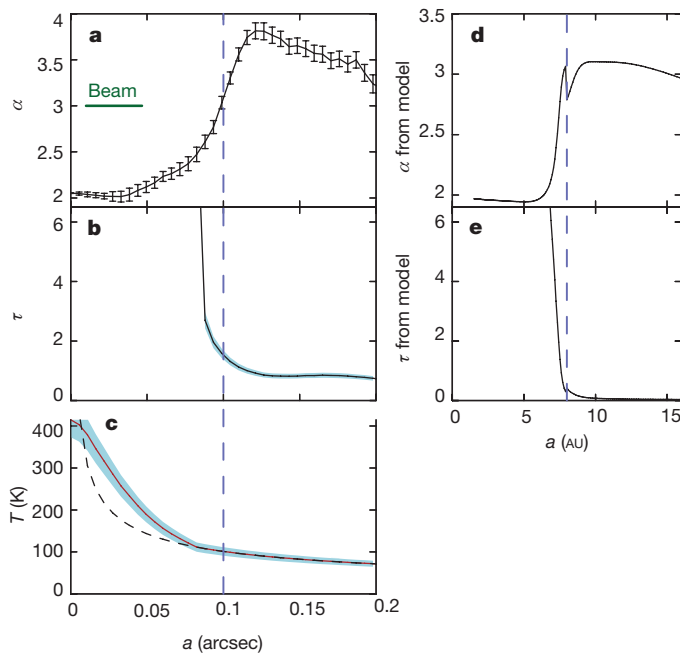


Figure 2 | Comparison of observations to models. **a–c**, The spectral index (α), the optical depth (τ), and the temperature (T) that we derived from the observed data, all as a function of the semi-major axis, a . The temperature profile, shown in red in **c**, is fixed to a square-root law (shown by the black dashed line) at the radii at which τ drops below 3. The uncertainties (error bars and light blue regions) are 68% confidence intervals (see Methods). The blue dashed line corresponds to the location of the water snow-line. α and τ are dimensionless. **d, e**, Predictions from the model in ref. 14 for a disk viscosity of 10^{-4} and an optically thick inner disk.

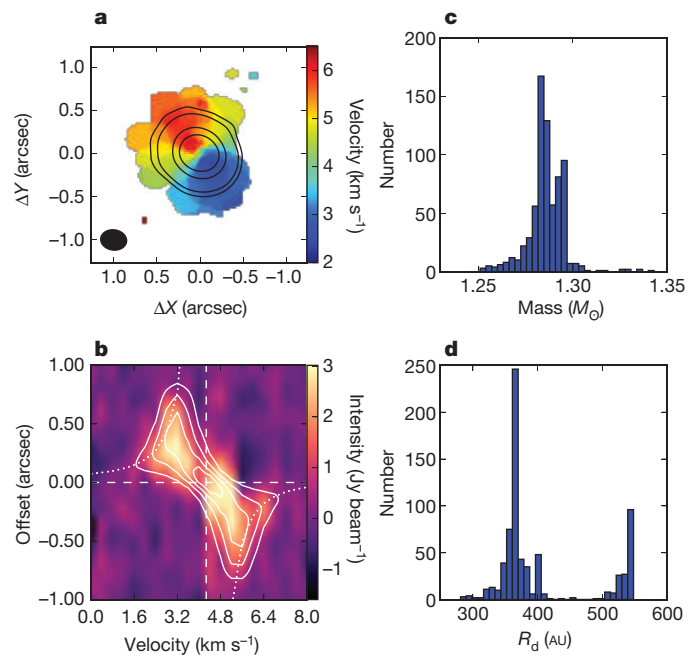


Figure 3 | Dynamical mass estimate. **a**, Intensity-weighted mean velocity map of the C^{18}O line, revealing clear Keplerian rotation. Continuum flux contours have been superposed. X and Y are the right ascension and the declination, respectively. **b**, Position–velocity diagram of the C^{18}O $J=2-1$ line. The horizontal and vertical dashed lines denote the position of the source and the source velocity. The curved dotted lines correspond to a Keplerian rotation curve, assuming a central mass of $1.3M_{\odot}$. **c, d**, Distribution of **c**, the masses of the central object, and **d**, the radii of the disk (R_d), resulting from the minimum χ^2 fittings.

out to $\sim 0.3''$ (125 AU). The brightness profile (Fig. 1b) indicates that what looks like a ring at $0.1''$ in the continuum image is really a sharp transition between these two regions.

Our continuum observations include two different spectral windows, each 1.875 GHz wide and centred at 218.0 GHz and 232.6 GHz, respectively (see Methods). Even though these spectral windows are separated by only 14.6 GHz, the very high signal-to-noise ratio of our observations allows us to derive accurate information regarding the spectral behaviour of the spatially resolved disk emission out to a radius of about $0.2''$ (85 AU). We use concentric ellipses (inclination = 38.3° and position angle = 32.4°) to extract radial profiles in each spectral window as a function of the semi-major axis, a . We compare the radial profiles extracted in the two spectral windows separately. In order to have perfectly matched UV coverage in the two spectral windows, we degraded the resolution of the 232.6 GHz observations to match that of the 218.0 GHz data. The spectral index, $\alpha = \ln(F_{232.6 \text{ GHz}}/F_{218.0 \text{ GHz}})/\ln(218.0 \text{ GHz}/232.6 \text{ GHz})$, as a function of a is shown in Fig. 2a (F is the flux at the frequency shown). We find distinct spectral behaviours across the disk, with $\alpha = 2.02 \pm 0.03$ in the central beam of the inner disk (corresponding to optically thick black-body emission, with temperatures of more than 100 K), and α reaching 3.7 ± 0.2 in the outer disk (typical of optically thin interstellar-medium values).

The observed spectral trends can be cast in terms of physical conditions, with grey-body fits²⁰ that can be used as diagnostics for the optical depth, $\tau(a) = \tau_0 \times (\nu/\nu_0)^\beta$, and for the average dust temperature along the line of sight (T_s , summed to $\tau \approx 1$) (where ν is the frequency of the emission and β is the power-law index relating τ and ν). In our case, the spectral information available is an amplitude and a slope, at the reference frequency $\nu_0 = 218.0$ GHz. Because we are provided with only two data points, we fix $\beta = 1.0$, as appropriate for circumstellar material²¹. The inner disk is very optically thick, and we can obtain an accurate estimate for T_s that is independent of the adopted β . On the other hand, τ_0 and T_s become degenerate in the optically thin regime.

We therefore adopt a temperature profile with $T_s \propto 1/\sqrt{a}$ extrapolated from the region where $\tau = 3$. The corresponding τ and T_s profiles for these assumptions are shown in Fig. 2b, c. We find that the sharp (unresolved) break at $\sim 0.1''$ seen in the V883 Ori disk (Fig. 1a) is associated with a steep drop in optical depth and with the transition from the optically thick to the optically thin regime. This result is robust and insensitive to β and to the exact prescription used to estimate τ and T_s beyond $0.1''$.

This intensity break occurs where the temperature has dropped below 105 ± 11 K. This temperature is more consistent with a water snow-line than with the snow-line of any of the other major volatiles present in protoplanetary disks (carbon monoxide, carbon dioxide and methane). The sublimation temperature of water is a strong function of ambient pressure. While it can be close to ~ 100 K in the interstellar medium, high-vacuum laboratory experiments and simulations^{22–24} suggest that it should be ~ 150 – 170 K at the 10^{-4} bar pressures expected at the location of the water snow-line in a typical disk (1 – 5 AU)^{11,25}. However, because the pressure is lower at ~ 40 AU, the sublimation temperature should also be lower in the case of V883 Ori. Furthermore, our temperature estimate is based on an extrapolation from the surface of the optically thick inner disk and might underestimate the true temperature of the water snow-line owing to the intense viscous heating at the disk midplane²⁵.

The observed spectral behaviour across the water snow-line has recently been predicted¹⁴ from numerical models that include the radial drift, coagulation, and fragmentation of dust grains. In Fig. 2 we show predictions for these models based on low disk viscosity ($\alpha_{\text{vis}} = 10^{-4}$), which result in an optically thick inner disk, as appropriate for V883 Ori. The model predictions are not convolved with the ALMA beam and thus have higher resolution than our observations. In these models, the fragmentation velocity of dust is 1 metre per second inside the snow-line and 10 metres per second for ice-covered grains outward of this line. In this scenario, icy grains quickly grow into centimetre-sized pebbles. Some of these icy particles drift into the inner disk, where their icy mantles evaporate. When this happens, their drift velocity

decreases, while the fragmentation efficiency increases. This produces an accumulation of millimetre-sized grains in the inner disk, driving the 230 GHz opacity up and the spectral index to the optically thick limit of ~ 2 . In these models¹⁴, α increases and τ decreases steeply around the water snow-line, in remarkably good agreement with our observational results. Our ALMA observations thus represent both a confirmation of the predictions of ref. 14 and the first resolved image of the signatures of the water snow-line in a protoplanetary disk.

By fitting a Keplerian model to the C¹⁸O line data, we derive a dynamical mass of $(1.3 \pm 0.1)M_{\odot}$ for the central source (see Fig. 3 and Methods). Assuming an age of 0.5 million years (Myr), as appropriate for a class I protostar such as V883 Ori (ref. 15), its photospheric luminosity should be a mere $\sim 6L_{\odot}$ (ref. 26). On the basis of the stellar mass and the observed luminosity of $400L_{\odot}$ (ref. 19), we derive an accretion rate of $7 \times 10^{-5}M_{\odot}$ per year, which is typical of FU Ori objects¹³. The location of the water snow-line in a protoplanetary disk is mostly determined by accretion heating in young Solar-type stars^{11,25}. For a $1M_{\odot}$ star, the snow-line begins at ~ 5 au at disk formation and moves inward to ~ 1 au by an age of a few million years, driven by the steady decrease in the accretion rate during disk evolution¹¹. However, as shown by V883 Ori, this steady evolution is punctuated by extreme bursts of accretion that can drive the snow-line out to more than 40 au.

In contrast to the HL Tau protoplanetary disk²⁷, whose concentric gaps have been interpreted as revealing the occurrence of planet formation at condensation fronts¹, the optical depth structure in V883 Ori is close to a step-function, as would be expected for efficient grain growth beyond a critical radius. Outward of the water snow-line, grains are covered by ice and can coagulate more efficiently into snowballs and eventually icy planetesimals²⁸. Inside the snow-line, on the other hand, ice mantles evaporate, increasing the efficiency of destructive collisions and resulting in the production of a new population of small dust grains²⁹. In this scenario, illustrated in Extended Data Fig. 1, an FU Ori outburst can increase the optical depth at millimetre wavelengths of a large region of the disk, by melting snowballs and releasing silicate grains from their icy mantles, in turn triggering further dust production. If the HL Tau ring system is in fact due to planet formation promoted by the condensation fronts, then the case of V883 Ori would represent an even earlier stage of disk evolution. Substantial evolution of solids (their growth, migration and fragmentation) has already occurred, but dynamical clearing of gaps by a planet has not yet happened.

While the fact that V883 Ori might reveal some of the very early steps towards planet formation is fascinating in itself, the outward movement of the water snow-line during FU Ori outbursts has far-reaching consequences for our understanding of disk evolution and planet formation in general. The water snow-line establishes the basic architecture of planetary systems like our own: in our Solar System, rocky planets formed inward of this line in the protosolar nebula, while giants formed outside. However, the intimate relation between the position of the water snow-line and the evolution of the central star is not yet understood. Although present population-synthesis models for planets do consider a steady decrease in the accretion rate during disk evolution and the corresponding inward motion of the water snow-line at the planet-formation epoch^{11,25}, they do not take into account the dramatic effects that FU Ori outbursts have on the snow-line location during the class I stage. If most systems experience FU-Ori-type outbursts during their evolution, as proposed in the episodic accretion scenario^{13,15,16}, then this implies that highly dynamical snow-lines must be taken into consideration in planet-formation models.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 March; accepted 18 May 2016.

1. Zhang, K., Blake, G. A. & Bergin, E. A. Evidence of fast pebble growth near condensation fronts in the HL Tau protoplanetary disk. *Astrophys. J.* **806**, L7–L12 (2015).

2. Okuzumi, S., Tanaka, H., Kobayashi, H. & Wada, K. Rapid coagulation of porous dust aggregates outside the snow line: a pathway to successful icy planetesimal formation. *Astrophys. J.* **752**, 106–123 (2012).
3. Guidi, G. *et al.* Dust properties across the CO snowline in the HD 163296 disk from ALMA and VLA observations. *Astron. Astrophys.* **588**, A112–A123 (2016).
4. Baillié, K., Charnoz, S. & Pantin, E. Time evolution of snow regions and planet traps in an evolving protoplanetary disk. *Astron. Astrophys.* **577**, A65–A76 (2015).
5. Blum, J. & Wurm, G. The growth mechanisms of macroscopic bodies in protoplanetary disks. *Annu. Rev. Astron. Astrophys.* **46**, 21–56 (2008).
6. Zhang, K. *et al.* On the commonality of 10–30 au sized axisymmetric dust structures in protoplanetary disks. *Astrophys. J.* **818**, L16–L22 (2016).
7. Qi, C. *et al.* Imaging of the CO snow line in a solar nebula analog. *Science* **341**, 630–632 (2013).
8. Nomura, H. *et al.* ALMA observations of a gap and a ring in the protoplanetary disk around TW Hya. *Astrophys. J.* **819**, L7–L13 (2016).
9. Schwarz, K. *et al.* The radial distribution of H₂ and CO in TW Hya as revealed by resolved ALMA observations of CO isotopologues. *Astrophys. J.* **823**, 91S (2016).
10. Qi, C. *et al.* Chemical imaging of the CO snow line in the HD 163296 disk. *Astrophys. J.* **813**, 128 (2015).
11. Kennedy, G. & Kenyon, S. Planet formation around stars of various masses: the snow line and the frequency of giant planets. *Astrophys. J.* **673**, 502–512 (2008).
12. Morbidelli, A., Lambrechts, M., Jacobson, S. & Bitsch, B. The great dichotomy of the Solar System: small terrestrial embryos and massive giant planet cores. *Icarus* **258**, 418–429 (2015).
13. Audard, M. *et al.* in *Protostars and Planets VI* (eds Beuther, H., Klessen, R. S., Dullemond, C. P. & Henning, T.) 387–410 (Univ. Arizona Press, 2014).
14. Banzatti, A. *et al.* Direct imaging of the water snow line at the time of planet formation using two ALMA continuum bands. *Astrophys. J.* **815**, L15–L20 (2015).
15. Evans, N. *et al.* The Spitzer c2d legacy results: star-formation fates and efficiencies; evolution and lifetimes. *Astrophys. J.* **181** (Suppl.), 321–350 (2009).
16. Dunham, M. & Vorobyov, E. Resolving the luminosity problem in low-mass star formation. *Astrophys. J.* **747**, 52–72 (2012).
17. Strom, K. & Strom, S. The discovery of two FU Orionis objects in L1641. *Astrophys. J.* **412**, L63–L66 (1993).
18. Menten, K. M., Reid, M. J., Forbrich, J. & Brunthaler, A. The distance to the Orion Nebula. *Astron. Astrophys.* **474**, 515–520 (2007).
19. Sandell, G. & Weintraub, D. On the similarity of FU Orionis stars to class I protostars: evidence from the submillimeter. *Astrophys. J.* **134** (Suppl.), 115–132 (2001).
20. Casassus, S. *et al.* A compact concentration of large grains in the HD 142527 protoplanetary dust trap. *Astrophys. J.* **812**, 126–139 (2015).
21. Williams, J. & Cieza, L. Protoplanetary disks and their evolution. *Annu. Rev. Astron. Astrophys.* **49**, 67–117 (2011).
22. Collings, M. *et al.* A laboratory survey of the thermal desorption of astrophysically relevant molecules. *Mon. Not. R. Astron. Soc.* **354**, 1133–1140 (2004).
23. Fayolle, E. *et al.* Laboratory H₂O:CO₂ ice desorption data: entrapment dependencies and its parameterization with an extended three-phase model. *Astron. Astrophys.* **529**, A74–A84 (2011).
24. Martín-Domènech, R., Muñoz Caro, G. M., Bueno, J. & Goesmann, F. Thermal desorption of circumstellar and cometary ice analogs. *Astron. Astrophys.* **564**, A8–A19 (2014).
25. Mulders, G., Ciesla, F., Min, M. & Pascucci, I. The snow line in viscous disks around low-mass stars: implications for water delivery to terrestrial planets in the habitable zone. *Astrophys. J.* **807**, 9–15 (2015).
26. Siess, L., Dufour, E. & Forestini, M. An internet server for pre-main sequence tracks of low- and intermediate-mass stars. *Astron. Astrophys.* **358**, 593–599 (2000).
27. ALMA Partnership *et al.* The 2014 ALMA long baseline campaign: first results from high angular resolution observations toward the HL Tau region. *Astrophys. J.* **808**, L3–L12 (2015).
28. Ros, K. & Johansen, A. Ice condensation as a planet formation mechanism. *Astron. Astrophys.* **552**, A137–A150 (2013).
29. Birnstiel, T., Dullemond, C. P. & Brauer, F. Gas- and dust evolution in protoplanetary disks. *Astron. Astrophys.* **513**, A79–A99 (2010).

Acknowledgements We thank the referees for their valuable comments. We also thank A. Banzatti and P. Pinilla for providing their model predictions in tabular form (Fig. 2d, e). ALMA is a partnership of the European Southern Observatory (ESO; representing its member states), the National Science Foundation (NSF; USA) and the National Institutes of Natural Sciences (Japan), together with the National Research Council (Canada) and the National Science Council and the Academia Sinica Institute of Astronomy and Astrophysics (Taiwan), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, Associated Universities Inc./National Radio Astronomy Observatory (NRAO), and the National Astronomical Observatory of Japan. The NRAO is a facility of the NSF, operated under cooperative agreement by Associated Universities. Support for this work was provided by the Millennium Science Initiative (Chilean Ministry of Economy), through grants RC130007 and IC120009. L.A.C., D.A.P., J.L.P. and C.C. acknowledge support from CONICYT

FONDECYT grants 1140109, 3150550, 1151445 and 3140592, respectively. H.C. acknowledges support from the Spanish Ministerio de Economía y Competitividad under grant AYA2014-55840P. Our work made use of ALMA data available at <https://almascience.eso.org/alma-data> with the following accession numbers: 2013.1.00710.S and 2015.1.00350.S.

Author Contributions L.A.C. led the ALMA cycle-2 and cycle-3 proposals (with the contribution of most of the other authors) and the writing of the manuscript. S.C. analysed the cycle-3 data and performed the grey-body analysis. J.T. and S.A.B. determined the stellar dynamical mass. J.P.W. analysed the cycle-2 molecular line data. S.P. and Z.Z. performed the simulations supporting the

cycle-3 proposal. All co-authors commented on the manuscript and contributed to the interpretation of the results.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.A.C. (lucas.cieza@mail.udp.cl).

Reviewer Information *Nature* thanks E. Bergin and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Our band-6 cycle-2 observations were taken under ALMA program 2013.1.00710.S with three different antenna configurations. V883 Ori was observed on 12 December 2014 and 5 April 2015, using 37 and 39 antennas on the C34-2/1 and C34-1/2 configurations, respectively. The minimum and maximum baselines of both configurations are very similar: minimum ~ 14 metres and maximum ~ 350 metres. The integration time was ~ 2 minutes in each epoch. The target was re-observed on 30 August 2015, with 35 antennas on the C34-7/6 configuration and with baselines in the range 42 metres to 1.5 kilometres. In this configuration, the integration time was ~ 3 minutes. For all antenna configurations, the ALMA correlator was configured so that three spectral windows with 58.6 MHz bandwidths were centred at 230.5380 GHz, 220.3987 GHz and 219.5603 GHz, to cover the $^{12}\text{CO } J=2-1$, $^{13}\text{CO } J=2-1$ and $\text{C}^{18}\text{O } J=2-1$ transitions, respectively. Two additional spectral windows with 1.875 GHz bandwidths were centred at 232.6 GHz and 218.0 GHz for continuum observations. The moon Ganymede and the quasar J0423-013 were used as flux calibrators, while the quasars J0538-4405 and J0607-0834 were observed for bandpass calibration. Observations of nearby phase calibrators (the quasars J0541-0541, J0532-0307 and/or J0529-0519) were alternated with our present target to calibrate the time-dependence variations of the complex gains.

V883 Ori was also observed in cycle-3 under program 2015.1.00350.S on 27 October 2015, with 45 antennas in the C38-8 configuration. This is the most extended array configuration offered in cycle-3, with baselines ranging from 267 metres to 12.6 kilometres. The total on-source integration time was 23 minutes. The correlator set-up was identical to that for our cycle-2 observations. J0541-0541 and J0529-0519 were used as primary and secondary phase calibrators, respectively. J0423-0120 was observed as a bandpass calibrator, and also as the primary flux calibrator. All data were calibrated using the Common Astronomy Software Applications package (CASA v4.4.0)³⁰ by the ALMA observatory. The standard calibration included offline water vapour radiometer (WVR) calibration, system temperature correction, and bandpass, phase and amplitude calibrations. Continuum images and spectral-line data cubes were created from the pipeline-calibrated visibilities using the CLEAN routine and Briggs weighting in the CASA v4.4.0 software package. Continuum subtraction was performed in the visibility domain before imaging the CO lines. Similarly, CLEANing of the dust continuum was performed after removing channels containing line emission.

All of the cycle-2 observations (three epochs with three different array configurations) were combined together to produce a single C^{18}O data cube. The root mean squared (r.m.s.) in this data cube is 10 mJy beam^{-1} per 0.25 km s^{-1} channel, with a beam of size $\sim 0.35''$ by $\sim 0.27''$ and a position angle of 89.9° . The long-baseline cycle-3 data set was reduced by itself in a similar fashion to the cycle-2 observations. The continuum data resulted in a $0.029'' \times 0.038''$ beam, and a r.m.s. of $0.05 \text{ mJy beam}^{-1}$, after one iteration of phase-only self-calibration.

The grey-body diagnostics used as a proxy for physical conditions of the dust require comparable UV coverages in both continuum frequencies. However, the difference in frequency in simultaneous observations implies a corresponding radial shift in the UV coverage. We followed two independent approaches to build such comparable maps, and confirmed that the two approaches provide very similar spectral-index maps and physical conditions. We first obtained two restored maps with multiscale Cotton-Schwab CLEAN³¹ by splitting off each spectral window, with the same CLEAN masks. We then degraded the higher frequency with an elliptical Gaussian whose axes correspond to the difference in quadrature of both clean beams. We also followed a second method, based on non-parametric Bayesian image synthesis. We fit an image model to the observations at each spectral window, and in the visibility domain, by minimizing the weighted least-square distance, as previously done in other multifrequency analyses of ALMA data^{20,32}. Because both approaches provided very similar trends, we adopted the Bayesian image synthesis, as it potentially allows for slightly finer angular detail, and the residual were more homogeneous (that is, free from structure) across the image than were the CLEAN residuals. Following ref. 20, we performed grey-body fits to our 218.0 GHz and 232.6 GHz images, such that $I_\nu(a) = B_\nu(T_s(a))[1 - \exp(-\tau(a))]$, where the optical depth is calculated as $\tau(a) = \tau_0 \times (\nu/\nu_0)^3$, I_ν is the specific intensity, B_ν is the spectral radiance of the grey body, and T_s is the average dust temperature along the line of sight (summed to $\tau \sim 1$). In our case, the spectral information available is an amplitude and a slope, at the reference frequency $\nu_0 = 218.0 \text{ GHz}$. We fixed $\beta = 1.0$, as appropriate for circumstellar material¹⁴. We estimated the error bars (68% confidence intervals) on the spectral indices from the r.m.s. scatter of specific intensities within each elliptical bin, summed in quadrature with the

r.m.s. intensity of the image-synthesis residuals. The uncertainties in the τ and temperature profiles (light blue regions in Fig. 2b, c) are given by a systematic flux calibration error of 10% (68% confidence interval).

Molecular-line kinematics around a central object are most commonly analysed by creating a position-velocity diagram. The C^{18}O line signal was too weak to be detected at the high spatial resolution of the cycle-3 observations; therefore, we use the cycle-2 data at $0.2''$ resolution to investigate the gas dynamics. Figure 3 shows the flux as a function of the velocity and position along the major axis of the disk. In this diagram, the position is shown as an offset from the centre of the disk, and the flux has been integrated along the width of the cut ($1.5''$ in the direction of the semi-major axis). The disk position angle was set to $32.4^\circ \pm 2^\circ$, determined from an elliptical Gaussian fit. The diagram shows separated blue-shifted and red-shifted components, suggesting Keplerian rotation. The radius of the disk visible in the diagram is $0.75''$, or about 320 AU . The central part around the source velocity (4.3 km s^{-1}) traces the outer slowly rotating material, and is largely resolved out owing to the extended emission. We also see that the data trace the higher velocities to 1.6 km s^{-1} and 6.6 km s^{-1} .

To further analyse the C^{18}O line emission and to give a mass estimate for the central object, we fitted a geometrically thin disk model, based on the model in ref. 33. On the basis of the channel maps and the position-velocity diagram, we decided to model a pure Keplerian disk without any infall. The velocity structure is then given by:

$$v_\phi(r) = \sqrt{\frac{GM}{r}}$$

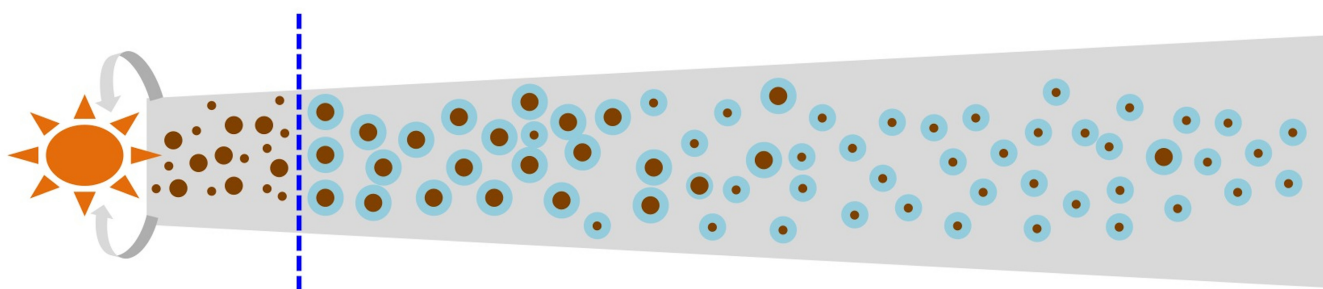
where v_ϕ is the velocity of the gas in the disk at location r , G is the gravitational constant, and M is the mass of the central protostar. We conducted a χ^2 -minimization fitting using the method of ref. 34:

$$\chi^2 = \frac{1}{N} \sum_{i=1}^N \left(\frac{D_i - M_i}{\sigma_{\text{r.m.s.}}} \right)^2$$

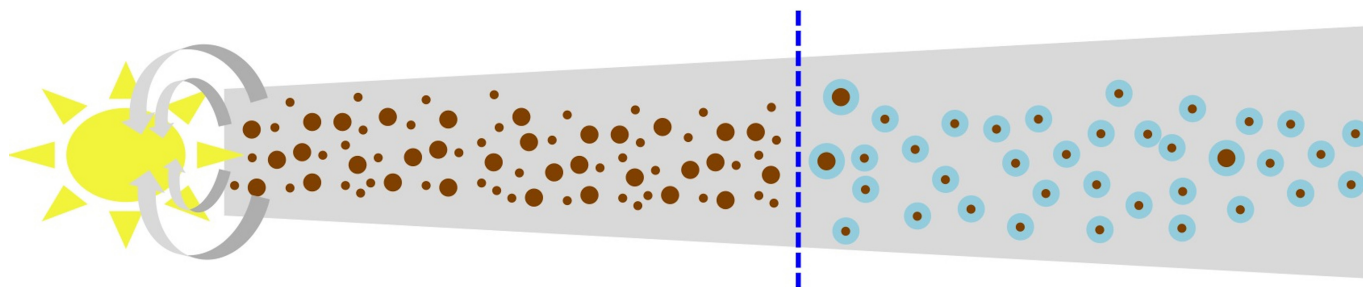
where D_i refers to the data, M_i the model and $\sigma_{\text{r.m.s.}}$ the observed noise error observed for every velocity channel map. As this method will not search through the entire parameter space, we carried out multiple runs using different random starting parameters. The parameters used by the fitting algorithm were: the mass of the central object (M), the radius of the Keplerian disk (R_d), the peak intensity (I_0) and the full-width at half-maximum (FWHM) of the Gaussian intensity distribution. The position angle was again set to 32.4° , and we determined an inclination of $38.3 \pm 1^\circ$ from the ratio of the major and minor axes of the disk in the continuum image. We adopted a distance of 414 parsecs (ref. 18), and set the source velocity (v_{src}) to 4.3 km s^{-1} on the basis of the moment-1 map. We fixed the centre of the disk to the centre of the continuum image. Of all these runs, the best fit (lowest χ^2) gave the following parameters: $M = (1.29 \pm 0.02)M_\odot$, $R_d = 361 \pm 27 \text{ AU}$, $I_0 = 0.18 \pm 0.04 \text{ Jy beam}^{-1}$ and $\text{FWHM} = 1.14'' \pm 0.11''$. The distributions of the parameters M and R_d are shown in Fig. 3. A mass of around $1.3M_\odot$ is the preferred solution. We adopt a 10% total error (68% confidence interval) in the dynamical mass to account for the uncertainty in the distance to V883 Ori, which dominates the error budget. There is more degeneracy in the size of the Keplerian disk, with values mainly varying between 300 AU and 550 AU. Using the parameters given by the best fit, we overlaid contours of the model over the C^{18}O emission as shown in Fig. 3. The shapes of the contours follow the data well. We see that the model traces higher velocities, which is to be expected as the model does not suffer from noise. A Keplerian rotation curve assuming a central mass of $1.3M_\odot$ is also plotted in Fig. 3. Most of the model and C^{18}O emission fall well within this curve.

30. McMullin, J. P. *et al.* CASA architecture and applications. *Astron. Soc. Pac. Conf. Ser.* **376**, 127–130 (2007).
31. Rau, U. & Cornwell, T. A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry. *Astron. Astrophys.* **532**, A71–A87 (2011).
32. Casassus, S. *et al.* Flows of gas through a protoplanetary gap. *Nature* **493**, 191–194 (2013).
33. Maret, S. Thindisk 1.0: compute the line emission from a geometrically thin protoplanetary disk. *Zenodo* <https://zenodo.org/record/13823> (2015).
34. Powell, M. J. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Computer J.* **7**, 155–162 (1964).

(a) Quiescent snow-line



(b) Outburst snow-line

**Extended Data Figure 1 | Sketch of the observed phenomenon.**

a, During quiescence, the water snow-line around stars of Solar masses is located 5 AU or less from the star, where the temperature of the disk reaches the sublimation point of water. **b**, During protostellar accretion outbursts, this line moves out to more than 40 AU, where it can be detected. Outward of the snow-line, grain growth is promoted by the high

coagulation efficiency of ice-covered grains (brown and blue concentric circles). Inward of this line, dust production is promoted by the high fragmentation efficiency of bared silicates (brown circles). This results in the observed break in the disk intensity profile, a steep reduction in the 1.3-mm dust opacity, and a sharp increase in the spectral index across the snow-line.

A sensitive electrometer based on a Rydberg atom in a Schrödinger–cat state

Adrien Facon¹, Eva–Katharina Dietsche¹, Dorian Grosso¹, Serge Haroche¹, Jean–Michel Raimond¹, Michel Brune¹ & Sébastien Gleyzes¹

Fundamental quantum fluctuations caused by the Heisenberg principle limit measurement precision¹. If the uncertainty is distributed equally between conjugate variables of the meter system, the measurement precision cannot exceed the standard quantum limit. When the meter is a large angular momentum, going beyond the standard quantum limit requires non-classical states such as squeezed states^{2–4} or Schrödinger–cat-like states^{5–7}. However, the metrological use of the latter^{8–10} has been so far restricted to meters with a relatively small total angular momentum because the experimental preparation of these non-classical states is very challenging^{11,12}. Here we report a measurement of an electric field based on an electrometer consisting of a large angular momentum (quantum number $J \approx 25$) carried by a single atom in a high-energy Rydberg state. We show that the fundamental Heisenberg limit¹³ can be approached when the Rydberg atom undergoes a non-classical evolution through Schrödinger–cat states. Using this method, we reach a single-shot sensitivity of 1.2 millivolts per centimetre for a 100-nanosecond interaction time, corresponding to 30 microvolts per centimetre per square root hertz at our 3 kilohertz repetition rate. This highly sensitive, non-invasive space- and time-resolved field measurement extends the realm of electrometric techniques^{14–17} and could have important practical applications: detection of individual electrons in mesoscopic devices^{18–21} at a distance of about 100 micrometres with a megahertz bandwidth is within reach.

Quantum metrology aims at measuring a classical quantity A (such as a frequency or a field) using a meter system (such as a harmonic oscillator or an angular momentum)^{2–4,7–10,22,23}, whose evolution depends upon A . The meter is initially prepared in a reference state and, after some interrogation time τ , its state is measured. The precision on the determination of A is limited by the intrinsic quantum fluctuations due to the probabilistic nature of the measurement. It can be traced to the properties of the meter's initial state, for which the uncertainty relation sets a lower bound to the product of uncertainty of non-commuting observables¹. Traditional schemes make use of classical states of the meter, the coherent states, where the uncertainty is distributed equally between conjugate observables²⁴. The precision of such scheme cannot exceed the standard quantum limit (SQL), which is directly derived from the Heisenberg inequality.

If the meter is a large spin J (which can be either the spin of a single particle, or the effective spin $J = N/2$ describing N two-level atoms) whose rotation frequency depends on A , the SQL derives from the non-commutation of the Cartesian spin components²⁵, and the smallest uncertainty achievable using spin coherent states (SCS)²⁶ scales as $1/\sqrt{J}$. To go beyond the SQL, the spin must be prepared in a non-classical state (note that in the case of an ensemble of two-level atoms, this requires the entanglement of the N systems¹³). Using spin squeezed states²⁵, the Heisenberg limit (HL)¹³, scaling as $1/J$, can be approached^{3,4}. It can even be reached with Schrödinger–cat states^{8–10}.

However, such states are difficult to prepare^{5,6,11,12}, and metrology experiments with cat states have been restricted so far to $J \approx 5$.

Here we report a cat-state-based measurement using a meter made of a large spin $J \approx 25$ carried by a single atom. We show that the HL can be approached when the spin undergoes a non-classical evolution through Schrödinger–cat states. Our experiment measures the amplitude F of an electric field F oriented along the O – z quantization axis (Extended Data Fig. 1). The J -spin system belongs to the Rydberg manifold $n = 50$ of a rubidium atom. Rydberg atoms have a very large polarizability, which makes them particularly suitable for measurements of small electric fields^{27,28}. The Stark levels in the manifold can be sorted by their magnetic quantum number m (Fig. 1a). We use the ladder made up of the lowest energy levels for each m , equidistant from each other to first order in F . This ladder is equivalent to that of the $|J, M\rangle$ levels of a spin $J = (n - 1)/2 = 49/2$ with $M = m - J$ (Methods). The spin levels have an energy $(M + J)\hbar\omega(F)$, with a spin angular frequency given by $\omega(F) = 3nFea_0/2\hbar$ (here a_0 is the Bohr radius), corresponding to $\partial\omega(F)/\partial F = 2\pi \times 96 \text{ MHz V}^{-1} \text{ cm}$. They are connected by σ_+ -polarized radio-frequency transitions. This spin evolves on a generalized Bloch sphere \mathcal{B} , $|J, J\rangle$ being the circular (C) Rydberg state nC at the north pole of \mathcal{B} . The spin coherent states²⁶, $|\theta, \varphi\rangle$, corresponding to a Bloch vector pointing in the θ, φ direction on \mathcal{B} , are defined by $|\theta, \varphi\rangle = \mathcal{R}(\theta, \varphi)|J, J\rangle$. The rotation operator $\mathcal{R}(\theta, \varphi)$ is realized by the application of a nearly resonant classical radio-frequency field with a Rabi frequency Ω_{rf} and a phase φ for a duration t such that $\theta = \Omega_{\text{rf}}t$.

A measurement at the SQL using SCS relies on a double radio-frequency pulse technique (Ramsey scheme, Fig. 1b). A first radio-frequency pulse at ω_{rf} prepares the $|\theta, 0\rangle$ SCS from the initial state $|J, J\rangle$. In a frame rotating at ω_{rf} around O – z , the further spin evolution is a precession at $\omega(F) - \omega_{\text{rf}}$, leading after an interrogation time τ to the $|\theta, \phi\rangle$ SCS with $\phi = (\omega(F) - \omega_{\text{rf}})\tau$. The field-sensitive phase ϕ is then read out by applying a final rotation $\mathcal{R}(\theta, \pi + \varphi_{\text{rf}})$ with an adjustable phase φ_{rf} . A measurement of the final spin state $|\psi_f\rangle$ provides information on F with a variance $\sigma_{F,\text{SQL}}^2$, where $\sigma_{F,\text{SQL}}^2$ is the single-shot SQL sensitivity:

$$\sigma_{F,\text{SQL}}^2 = \frac{1}{\tau \sqrt{2J}} \left| \frac{\partial \omega}{\partial F} \right|^{-1}. \quad (1)$$

In order to beat the SQL, we measure, instead of ϕ , the global quantum phase Φ accumulated by the spin J during its evolution on \mathcal{B} . Measuring Φ as a function of F requires a quantum reference state $|R\rangle$, unaffected by the spin J successive transformations. We use, for $|R\rangle$, the $n = 51$ circular state (Fig. 1a), which experiences no first order Stark shift. We initially prepare the superposition $1/\sqrt{2}(|J, J\rangle + |R\rangle)$ using a classical microwave pulse. The $|J, J\rangle$ part of this initial state then undergoes the Ramsey sequence sketched in Fig. 1b, ending in state $|\psi_f\rangle$, Φ being the phase of $\langle \psi_f | J | \psi_f \rangle$. We then apply a second $\pi/2$ microwave pulse, selectively

¹Laboratoire Kastler Brossel, Collège de France, CNRS, ENS-PSL Research University, UPMC-Sorbonne Universités, 11 place Marcelin Berthelot, 75231 Paris Cedex 05, France.

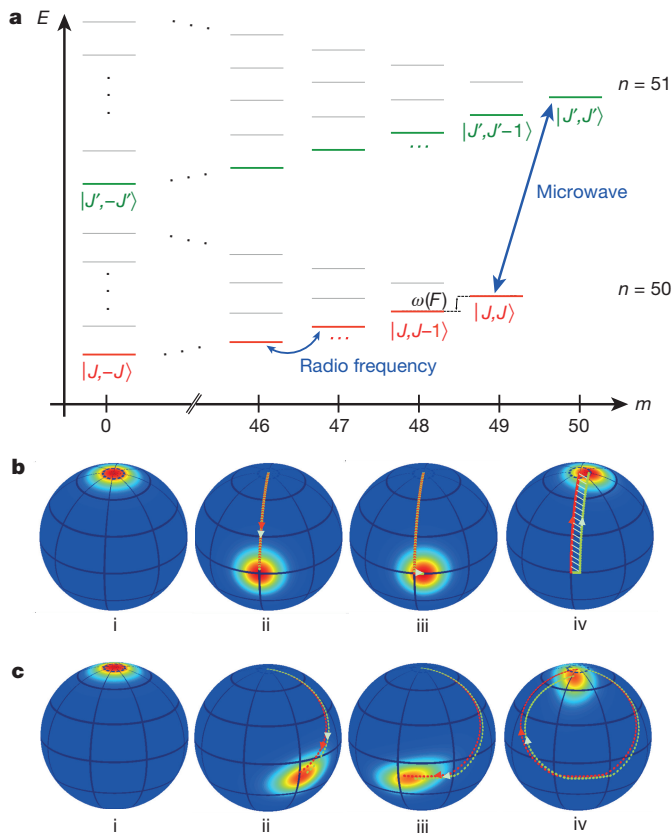


Figure 1 | Atomic levels and measurement sequence. **a**, Energies (E) of the Stark levels in the $n = 50$ and 51 manifold sorted by their magnetic quantum number, m (not to scale). The $m \geq 0$ levels of each manifold form a triangular structure with the circular state at its tip; this state is not displaced to first order by the electric field. The J spin states in the 50 manifold are depicted by the thick red lines. The reference state is the circular state $|J', J'\rangle$ in the 51 manifold belonging to the J' state ladder (green lines). **b**, Evolution of the J spin in the Ramsey sequence (Stark effect to first order in F). Shown are successive plots of the spin state Q -function in the rotating frame for $F = F_0 + \Delta F$. Initial $|J, J\rangle$ circular state (i). State after the first radio-frequency pulse inducing a $\theta = \pi/2$ rotation (ii). State after the interrogation time τ , before the second radio-frequency pulse (iii). Final state after the second radio-frequency pulse with $\varphi_{\text{rf}} = 0$ (iv). The green dotted line shows the trajectory on the Bloch sphere \mathcal{B} . The red line corresponds to the spin trajectory for $F = F_0$. The classical Ramsey scheme measures ΔF through the final position of the spin in iv, and is therefore limited by the quantum fluctuations of the SCS. In the cat-state-based scheme, we deduce ΔF from the global phase accumulated during the complete evolution, proportional to the dashed area. **c**, Simulation of a realistic sequence for $\tau = 56$ ns and $\Delta F = 1.7$ mV cm $^{-1}$, taking into account the second-order Stark effect and the finite duration of the radio-frequency pulses (184 ns). The phase φ_{rf} is chosen so that the trajectory is closed for $F = F_0$. The Q -function and the green line correspond to $F = F_0 + \Delta F$. The red line corresponds to the spin trajectory for $F = F_0$. The value of ΔF is deduced from the difference of the global phases accumulated along the red and green trajectories.

addressing the $|J, J\rangle \rightarrow |R\rangle$ transition, with an adjustable phase φ_{mw} . We finally measure whether the atom is in the state $|J, J\rangle$ or not (Methods).

The probability of finding the atom in $|J, J\rangle$ oscillates as a function of Φ . This oscillation is sensitive to small variations of F , since the atomic system is cast during the interrogation time τ in a quantum superposition of two states with different classical orbits, $|R\rangle$ and $|\theta, \varphi\rangle$ which have quite different electric dipoles (Methods), and is an atomic cat state²⁹. The interference phase Φ depends upon the exact spin trajectory on \mathcal{B} and thus upon F and φ_{rf} (Methods). The amplitude of the interference pattern is proportional to $|\langle \psi | J | J \rangle|$. It is maximum for $\varphi_{\text{rf}} \approx 0$ when the field F is close to the reference field F_0 such that

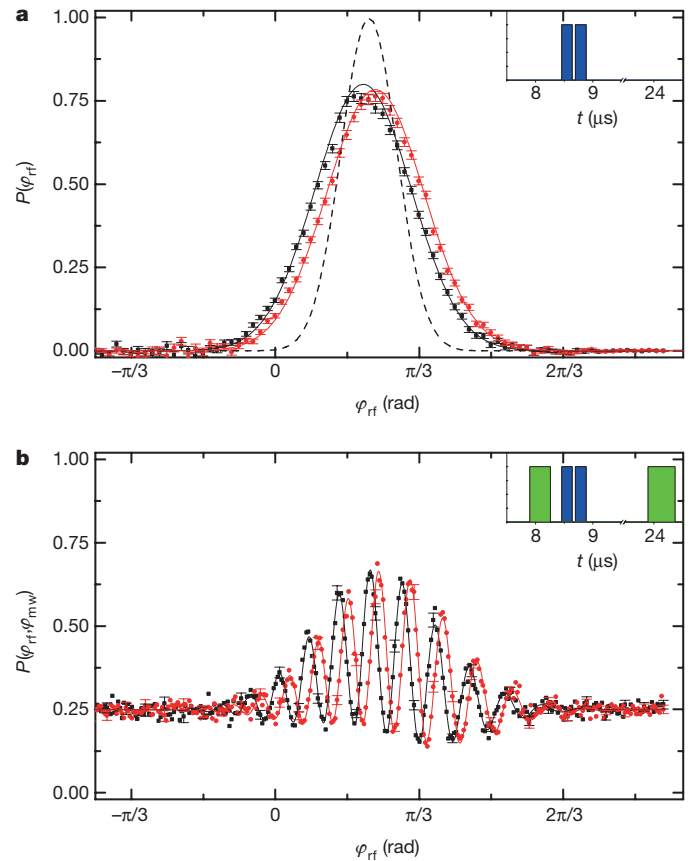


Figure 2 | Classical and cat-state-based field measurements. **a**, Simple Ramsey measurement. Shown is the probability $P(\varphi_{\text{rf}})$ of the spin ending up in $|J, J\rangle$. The radio-frequency pulses (blue bars in the timing inset), separated by the interrogation time $\tau = 56$ ns, have duration $t_2 = 184$ ns and correspond to $\Omega_{\text{rf}} t_2 = 1.86$ rad (Methods). The red and black experimental points correspond respectively to the field values $F_0 + \delta F/2$ and $F_0 - \delta F/2$, with $\delta F = 566$ $\mu\text{V cm}^{-1}$. The two signals are shifted by $\delta\phi = 82$ mrad. The reference phase φ_{rf}^0 is determined as the average of the centres of these curves. The error bars reflect the s.e.m. over 3,100 realizations of the experiment. The solid lines result from numerical simulations of the full experiment. The dashed line is a Gaussian with a width determined by the SQL. **b**, Results of 950 realizations of the same experiment including the microwave pulses. The timing is in the inset (green, microwave pulses; blue, radio-frequency pulses). The data points are experimental, with s.e.m. error bars every tenth point, with solid lines to guide the eye. The signals corresponding to $F_0 + \delta F/2$ and $F_0 - \delta F/2$ are shifted by about the same $\delta\phi$ as in **a**. The spacing between interference fringes being much smaller than the width of the Gaussian in **a**, the cat-state-based measurement is much more sensitive to variations of the electric field. The background probability for φ_{rf} far from φ_{rf}^0 is $1/4$. When the spin does not return close to the initial state, the $n = 50$ manifold does not contribute to the signal. However, half of the 50% population stored in $n = 51$ returns to $n = 50$ after the final microwave pulse.

$\omega(F_0) = \omega_{\text{rf}}$. Then, Φ can be expanded to first order in a small field variation $dF = F - F_0$ as:

$$\Phi \approx \Phi_0 + J (1 - \cos\theta) \left(\frac{\partial\omega}{\partial F} \right) dF\tau \quad (2)$$

where Φ_0 is the total phase accumulated for the reference field F_0 (Methods). This leads to a single-shot measurement sensitivity:

$$\sigma_F^1 = \frac{1}{\tau J (1 - \cos\theta)} \left| \frac{\partial\omega}{\partial F} \right|^{-1} \quad (3)$$

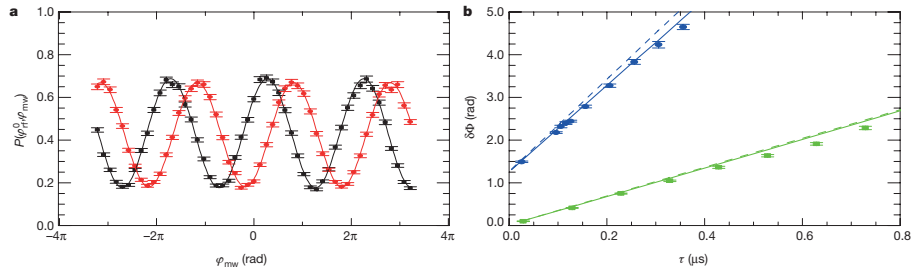


Figure 3 | Microwave Ramsey fringes. **a**, Probability $P(\varphi_{\text{rf}}^0, \varphi_{\text{mw}})$ of the spin being in $|J, J\rangle$ at the end of the complete sequence (experimental conditions are the same as for Fig. 2b), as a function of the relative phase φ_{mw} of the microwave pulses. The red and black dots correspond to experiments at $F_0 + \delta F/2$ and $F_0 - \delta F/2$, respectively (error bars are the s.e.m. over 3,100 realizations). The solid lines are sine fits, providing the relative phase $\delta\Phi$ and the contrast C of the interference signals. **b**, Phase increment $\delta\Phi$ corresponding to δF as a function of the interrogation

scaling as $1/J$. The factor $J(1 - \cos\theta)$, proportional to the difference of the electric dipoles of the components of the superposition, measures the ‘size’ of the Schrödinger-cat state. The Heisenberg limit, $\sigma_{F,\text{HL}}^1 = (1/2J\tau)|\partial\omega/\partial F|^{-1}$, is reached for $\theta = \pi$, when this size is maximum.

In the real experiment, we must take into account the finite duration of the radio-frequency pulses ($\Omega_{\text{rf}}/2\pi = 1.6$ MHz) and the second order Stark effect in the $n = 50$ manifold, which makes the spin states ladder slightly anharmonic. The trajectory of the spin on \mathcal{B} and the spin coherent states are distorted accordingly (Fig. 1c). The optimal phase for the second radio-frequency pulse is thus $\varphi_{\text{rf}} = \varphi_{\text{rf}}^0 \neq 0$. The state distortion slightly affects the contrast of the interferometric signal. Nevertheless, the main conclusions of the simple case discussion above remain valid.

Figure 2a presents, for reference, the results of the classical Ramsey method, in which no microwave pulses are applied (timing in the inset). We measure the probability $P(\varphi_{\text{rf}})$ of the spin returning to the $|J, J\rangle$ state as a function of φ_{rf} for two electric fields, $F_0 - \delta F/2$ and $F_0 + \delta F/2$, with $\delta F = 566 \mu\text{V cm}^{-1}$. These probabilities are Gaussian (Methods) centred around $\varphi_{\text{rf}}^0 = 0.715$ rad. The contrast is slightly reduced and the width increased with respect to the ideal case due to the second order Stark effect. The phase shift ($\delta\phi = 82$ mrad) induced by the variation δF of the electric field is small in comparison to the width of the signal (~ 0.722 rad).

Let us now consider the complete sequence (timing shown in the inset of Fig. 2b), with a fixed microwave phase $\varphi_{\text{mw}} = \varphi_{\text{mw}}^0$. We measure the probability $P(\varphi_{\text{rf}}, \varphi_{\text{mw}}^0)$ of finally detecting the atom in the initial $|J, J\rangle$ state as a function of φ_{rf} for the electric fields $F_0 - \delta F/2$ and $F_0 + \delta F/2$. This probability exhibits an interference pattern around $\varphi_{\text{rf}} = \varphi_{\text{rf}}^0$, revealing the rapid variation of Φ with φ_{rf} (Methods). The contrast of the interference reflects the probability amplitude of the spin J returning to its initial $|J, J\rangle$ state. Beyond the effect of the second order Stark shift, this contrast is further reduced by static electric field inhomogeneity, electric field noise and other experimental imperfections. This signal already demonstrates an improvement over the classical method. The measurement sensitivity is determined by the difference between the values of the probability P corresponding to the two electric fields for the same φ_{rf} . Although the horizontal shift between the red and black curves in Fig. 2a and Fig. 2b are comparable, the curves of Fig. 2b present much steeper features. Therefore, close to the centre of the interference pattern (Fig. 2b), the variation of P induced by a change δF is larger than the largest vertical distance between the red and black curves in Fig. 2a.

In order to assess the improvement over the SQL, we set $\varphi_{\text{rf}} = \varphi_{\text{rf}}^0$ and we record $P(\varphi_{\text{rf}}^0, \varphi_{\text{mw}})$ as a function of φ_{mw} for $F_0 - \delta F/2$ and $F_0 + \delta F/2$. Figure 3a presents the fringe signals together with sine fits. We extract from these fits the contrast C and the relative phase

time τ . The green and blue points (error bars resulting from the standard error of the fringe fits) correspond to the radio-frequency pulse durations t_1 and t_2 , respectively. The slope of the dashed lines correspond to the ideal model with no quadratic Stark effect. The solid lines are predictions of a numerical model including the second order Stark effect, which distorts the trajectory on \mathcal{B} and slightly reduces the value θ with respect to the ideal value $\Omega_{\text{rf}}t_i$ ($i = 1, 2$).

$\delta\Phi = 1.72$ rad of the two interference patterns, which is $21 \approx J$ times larger than the phase shift $\delta\phi$ obtained with the classical method (Methods). We have checked that $\delta\Phi$ is proportional to δF .

Figure 3b presents $\delta\Phi$ as a function of the interrogation time τ , for two radio-frequency pulse durations $t_1 = 91$ ns and $t_2 = 184$ ns and hence two θ values (Methods). We observe that $\delta\Phi$ grows linearly with τ , with a slope increasing with θ . The experimental data are in good agreement with the predictions of equation (2) (dashed lines). The agreement is improved by taking into account the second order Stark effect (solid lines). Note that for $\tau \rightarrow 0$, $\delta\Phi \rightarrow \delta\Phi_0 \neq 0$. This is due to the finite duration of the radio-frequency pulses, during which the spin state acquires a field-dependent phase along its path on \mathcal{B} .

To assess the measurement performance, we only consider the phase shift $\delta\Phi_\tau = \delta\Phi - \delta\Phi_0$ accumulated during the interrogation time τ . The single-shot sensitivity is then:

$$\sigma_F^1 = \frac{1}{C} \frac{\delta F}{\delta\Phi_\tau} \quad (4)$$

Figure 4 compares σ_F^1 to the SQL and HL as function of τ , for the t_2 radio-frequency pulse duration (blue points). For short interrogation times, the experimental points are well below the SQL. For larger τ values, the contrast C is reduced by experimental imperfections. In fact, this reduction is in part a direct consequence of the extreme sensitivity of the measurement. Electric field noise integrated over long times blurs the interference pattern.

The best single-shot sensitivity is $\sigma_F^1 = 1.2 \text{ mV cm}^{-1}$ for $\tau = 200$ ns. The experiment repetition rate is limited by the total sequence duration, 300 μs , dominated by the atomic time of flight from preparation to detection. The sensitivity is thus $30 \mu\text{V cm}^{-1} \text{ Hz}^{-0.5}$ corresponding to the possible detection, in 1 s, of a single electron at a 700 μm distance from the atom. This is, at least, an improvement of two orders of magnitude on the sensitivity reached by nitrogen-vacancy centres¹⁶ or quantum dots^{14,15,17}. Our experiment competes with the best electro-mechanical resonators^{18,19} or single-electron transistors (SET)^{20,21}, which provide sensitivities of the order of $10^{-6} \text{ e Hz}^{-0.5}$ at distances in the micrometre range, corresponding to $14 \mu\text{V cm}^{-1} \text{ Hz}^{-0.5}$. Note, furthermore, that the experimental sequence duration could easily be reduced down to a few microseconds by detecting the atoms in the interaction zone. The sensitivity would then reach an unprecedented $3 \mu\text{V cm}^{-1} \text{ Hz}^{-0.5}$.

We have shown that our method performs a very sensitive measurement of minute electric field variations, with a non-invasive probe made of a single Rydberg atom, considerably extending the metrologic applications of these states. The measurement time is short (in the ~ 100 ns range) making it possible to sample tiny variations of the electric field with a megahertz bandwidth. It could moreover be

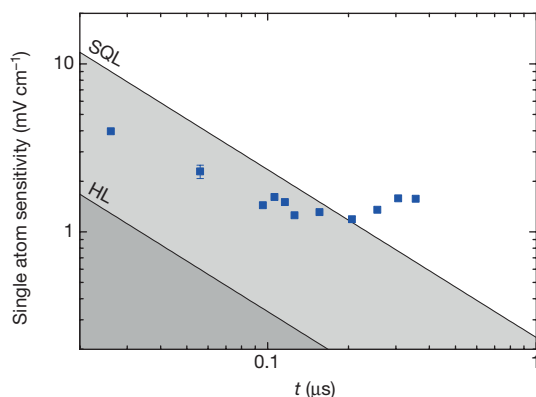


Figure 4 | Comparison of the single-shot sensitivity with the SQL and the HL. The blue points show a logarithmic plot of σ_F as a function of the interrogation time τ . The error bar on the point at $\tau = 56$ ns reflects the statistical s.d. of three experiments performed in the same conditions. The SQL and HL are depicted by black lines. The dark grey area is forbidden. The light-grey area corresponds to measurements beyond the SQL. The sensitivity of the experiment outperforms the SQL for $\tau < 200$ ns, by a factor of up to 2 for the shortest time.

resolved in space, with a few micrometres resolution, using Rydberg atoms excited in cold, trapped atom samples³⁰.

The sensitivity could be brought much closer to the HL by using an improved electrode design for a better field homogeneity and increased radio-frequency power making it possible to reach $\theta = \pi$ in spite of the second order Stark effect. An interrogation time of 200 ns would then correspond to $\sigma_F^1 = 170 \mu\text{V cm}^{-1}$. For a slightly longer 1 μs interrogation time, the phase-shift of the fringes for a field increment of $200 \mu\text{V cm}^{-1}$ (that of a single electron at a $270 \mu\text{m}$ distance) reaches π , allowing us in principle to distinguish two field values differing by this tiny amount with a single atomic detection (a few atoms, if $C < 1$).

This could lead to interesting applications in mesoscopic physics. The presence or absence of an electron in a quantum dot, realized in a 2D semiconductor or in a carbon nanotube, could be probed with a megahertz bandwidth by a few atoms located far away from the mesoscopic structure. Also, compared with SET detectors, this method does not have demanding cryogenic requirements, operates at large distances and does not require any modification of the device under test.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 20 January; accepted 5 May 2016.

- Giovannetti, V., Lloyd, S. & Maccone, L. Quantum-enhanced measurements: beating the standard quantum limit. *Science* **306**, 1330–1336 (2004).
- Wasilewski, W. *et al.* Quantum noise limited and entanglement-assisted magnetometry. *Phys. Rev. Lett.* **104**, 133601 (2010).
- Müssel, W., Strobel, H., Linnemann, D., Hume, D. B. & Oberthaler, M. K. Scalable spin squeezing for quantum-enhanced magnetometry with Bose-Einstein condensates. *Phys. Rev. Lett.* **113**, 103004 (2014).
- Hosten, O., Engels, N. J., Krishnakumar, R. & Kasevich, M. A. Measurement noise 100 times lower than the quantum-projection limit using entangled atoms. *Nature* **529**, 505–508 (2016).
- Massar, S. & Polzik, E. S. Generating a superposition of spin states in an atomic ensemble. *Phys. Rev. Lett.* **91**, 060401 (2003).

- Lau, H. W., Dutton, Z., Wang, T. & Simon, C. Proposal for the creation and optical detection of spin cat states in Bose-Einstein condensates. *Phys. Rev. Lett.* **113**, 090401 (2014).
- Tanaka, T. *et al.* Proposed robust entanglement-based magnetic field sensor beyond the standard quantum limit. *Phys. Rev. Lett.* **115**, 170801 (2015).
- Leibfried, D. *et al.* Toward Heisenberg-limited spectroscopy with multiparticle entangled states. *Science* **304**, 1476–1478 (2004).
- Nagata, T., Okamoto, R., O'Brien, J. L., Sasaki, K. & Takeuchi, S. Beating the standard quantum limit with four-entangled photons. *Science* **316**, 726–729 (2007).
- Jones, J. A. *et al.* Magnetic field sensing beyond the standard quantum limit using 10-spin NOON states. *Science* **324**, 1166–1168 (2009).
- Monz, T. *et al.* 14-Qubit entanglement: creation and coherence. *Phys. Rev. Lett.* **106**, 130506 (2011).
- Signoles, A. *et al.* Confined quantum Zeno dynamics of a watched atomic arrow. *Nat. Phys.* **10**, 715–719 (2014).
- Giovannetti, V., Lloyd, S. & Maccone, L. Advances in quantum metrology. *Nat. Photon.* **5**, 222–229 (2011).
- Vamivakas, A. N. *et al.* Nanoscale optical electrometer. *Phys. Rev. Lett.* **107**, 166802 (2011).
- Houel, J. *et al.* Probing single-charge fluctuations at a GaAs/AlAs interface using laser spectroscopy on a nearby InGaAs quantum dot. *Phys. Rev. Lett.* **108**, 107401 (2012).
- Dolde, F. *et al.* Nanoscale detection of a single fundamental charge in ambient conditions using the NV⁻ center in diamond. *Phys. Rev. Lett.* **112**, 097603 (2014).
- Arnold, C. *et al.* Cavity-enhanced real-time monitoring of single-charge jumps at the microsecond time scale. *Phys. Rev. X* **4**, 021004 (2014).
- Cleland, A. N. & Roukes, M. L. A nanometre-scale mechanical electrometer. *Nature* **392**, 160–162 (1998).
- Bunch, J. S. *et al.* Electromechanical resonators from graphene sheets. *Science* **315**, 490–493 (2007).
- Yoo, M. J. *et al.* Scanning single-electron transistor microscopy: imaging individual charges. *Science* **276**, 579–582 (1997).
- Devoret, M. H. & Schoelkopf, R. J. Amplifying quantum signals with the single-electron transistor. *Nature* **406**, 1039–1046 (2000).
- Lo, H. Y. *et al.* Spin-motion entanglement and state diagnosis with squeezed oscillator wavepackets. *Nature* **521**, 336–339 (2015).
- Wollman, E. E. *et al.* Quantum squeezing of motion in a mechanical resonator. *Science* **349**, 952–955 (2015).
- Caves, C. M., Thorne, K. S., Drever, R. W., Sandberg, V. D. & Zimmermann, M. On the measurement of a weak classical force coupled to a quantum-mechanical oscillator. I. Issues of principle. *Rev. Mod. Phys.* **52**, 341–392 (1980).
- Kitagawa, M. & Ueda, M. Squeezed spin states. *Phys. Rev. A* **47**, 5138–5143 (1993).
- Arecchi, F., Courtens, E., Gilmore, R. & Thomas, H. Atomic coherent spin states in quantum optics. *Phys. Rev. A* **6**, 2211–2237 (1972).
- Osterwalder, A. & Merkt, F. Using high Rydberg states as electric field sensors. *Phys. Rev. Lett.* **82**, 1831–1834 (1999).
- Abel, R. P., Carr, C., Krohn, U. & Adams, C. S. Electrometry near a dielectric surface using Rydberg electromagnetically induced transparency. *Phys. Rev. A* **84**, 023408 (2011).
- Hempel, C. *et al.* Entanglement-enhanced detection of single-photon scattering events. *Nat. Photon.* **7**, 630–633 (2013).
- Hermann-Avigilano, C. *et al.* Long coherence times for Rydberg qubits on a superconducting atom chip. *Phys. Rev. A* **90**, 040502(R) (2014).

Acknowledgements We thank A. Cottet, T. Kontos and W. Munro for discussions. We acknowledge funding by the EU under the ERC project ‘DECLIC’ and the RIA project ‘RYSQ’.

Author Contributions A.F., E.K.D., D.G., S.H., J.M.R., M.B. and S.G. contributed to the experimental set-up. A.F. and E.K.D. collected the data and analysed the results. J.M.R., S.H. and M.B. supervised the research. S.G. led the experiment. All authors discussed the results and the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.G. (gleyzes@lkb.ens.fr).

Reviewer Information *Nature* thanks C. Adams, L. Maccone and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Experimental set-up. The set-up (Extended Data Fig. 1) is made up of two parallel, horizontal disk electrodes, which create the vertical electric field F aligned along the O - z quantization axis. The gap between these electrodes is surrounded by four electrodes forming a ring, used to generate the rf field. The Rydberg atoms are excited stepwise by three laser beams at 780 nm, 776 nm and 1,259 nm resonant with the $5S_{1/2} \rightarrow 5P_{3/2}$, $5P_{3/2} \rightarrow 5D_{5/2}$ and $5D_{5/2} \rightarrow 49F$ transitions. They cross at a 45° angle the horizontal atomic beam at the centre of the electrode structure. The Doppler effect provides an atomic velocity selection at $v = 252 \pm 7$ m/s. The 780 and 776 nm cw laser beams are collinear, perpendicular to the third one. Every 311 μ s, a 0.5 μ s pulse of the 1,259 nm laser excites less than one rubidium atom on average into the $49F, m = 2$ state. This pulse sets the time origin $t = 0$ for each sequence. The quantization axis during the laser excitation is parallel to the 780 nm laser, and is defined by a dc field applied across the ring electrodes. This field is adiabatically switched off in 1 μ s, while a 2 V/cm field is switched on along O - z , which becomes the quantization axis during the measurement sequence. The atoms are then transferred in 2.7 μ s into the circular state using an adiabatic rapid passage¹² in a rf field at 230 MHz. The electric field is then ramped up in 0.5 μ s to $F = F_0 \pm \delta F$ with $F_0 = 5.50527 \pm 0.00021$ V/cm ($\omega(F_0)/2\pi = 530.019 \pm 0.020$ MHz). The state preparation sequence ends with a 0.5- μ s microwave pulse transferring $49C$ into $50C$. This excitation, selective in the magnetic quantum number m , ensures that spuriously prepared elliptical states remain in the $n = 49$ manifold and do not affect the experimental signals.

The σ^+ -polarized rf pulses are created by applying on two adjacent ring electrodes signals generated by 530 MHz synthesizers with finely tuned amplitudes and phases. The first pulse starts at $t = 8.5 \mu$ s. The two microwave $\pi/2$ pulses, starting at $t = 7.9 \mu$ s and $t = 23.9 \mu$ s, are generated from the same microwave source, a frequency-multiplied X-band synthesizer. They are tuned to 51.091 GHz, on resonance with the $50C \rightarrow 51C$ transition in the F_0 field.

After the end of the measurement, we measure the population of $|J, J\rangle$ by applying at $t = 30 \mu$ s a last m -selective microwave π -pulse tuned on the $50C$ – $52C$ two-photon transition and by detecting the $52C$ level by field ionization in the detector D.

Rf pulse duration optimization. The rf pulses at 530 MHz have a Rabi frequency $\Omega_{rf}/2\pi = 1.6$ MHz. The $51C$ circular reference state $|R\rangle$ is the $|J', J'\rangle$ state of a $J' = 25$ spin evolving on a Bloch sphere B' . Owing to the different Stark polarizabilities in the 50 and 51 manifolds, the rf field is 11 MHz out of resonance for the J' spin ladder and barely affects it. The rf pulse results only in a small J' spin precession at ~ 11 MHz near the north pole of B' . Moreover, we optimize its duration ($t_1 = 91$ ns or $t_2 = 184$ ns) so that the J' spin performs exactly one or two complete rotations, returning finally to its initial $|J', J'\rangle$ state. The corresponding θ values are $\theta_1 = \Omega_{rf}t_1 = 0.92$ rad and $\theta_2 = 1.86$ rad.

Calibration of the electric field. To measure the effect induced by a variation δF of the electric field, we alternate between an experimental sequence where we apply $F_0 - \delta F/2$ and a sequence where we apply $F_0 + \delta F/2$. F_0 is calibrated by measuring by standard spectroscopy the two transitions $|49C\rangle \rightarrow |J, J\rangle$ and $|49C\rangle \rightarrow |J, J-1\rangle$ separated by the frequency $\omega(F_0)/2\pi$. We found $\omega(F_0)/2\pi = 530.019 \pm 0.020$ MHz, corresponding to $F_0 = 5.50527 \pm 0.00021$ V/cm. The precision of the measurement is limited by the long term drift of the electric field (the error corresponds to the standard deviation over a few days of measurements). We also find $\delta F = 566 \pm 13$ μ V/cm, corresponding to $\delta\omega/2\pi = 54.8 \pm 1.2$ kHz.

Determination of the contrast C of the fringes. The long term electric field drift affects the contrast of the interference fringes. To get their intrinsic contrast leading to the sensitivity values displayed in Fig. 4, we alternate sequences with $F_0 + \delta F/2$ and $F_0 - \delta F/2$. We then use half of the data corresponding to $F_0 + \delta F/2$ to determine the slow phase drift of the interference fringes. This measured drift is used to post-process the other statistically independent half of the data, from which we deduce C .

Code availability. Requests for materials can be addressed to the corresponding author.

Analytic expression of Φ . To derive the expression of the phase Φ , we consider the evolution of J in the rotating frame at frequency ω_{rf} . We set the energy origin at that of the circular state $|50C\rangle = |J, J\rangle$. The first rf pulse induces a rotation $\mathcal{R}(\theta, 0)$, preparing $|\theta, 0\rangle = \mathcal{R}(\theta, 0)|J, J\rangle$. During the interrogation time τ , J rotates along the z axis of the Bloch sphere at a precession frequency $\delta\omega = \omega(F) - \omega_{rf}$, leading to the state $|\theta, \phi\rangle$, with $\phi = \delta\omega \tau$. Finally, a second rotation $\mathcal{R}(\theta, \pi + \varphi_{rf})$, brings the coherent spin state in the final state $|\psi_f\rangle$.

The phase Φ is defined from the overlap between $|\psi_f\rangle$ and $|J, J\rangle$:

$$\langle J, J | \psi_f \rangle = |\langle J, J | \psi_f \rangle| e^{-i\Phi}$$

Using $\langle J, J | \psi_f \rangle = \langle J, J | \mathcal{R}(\theta, \pi + \varphi_{rf}) | \theta, \phi \rangle = \langle \theta, \varphi_{rf} | \theta, \phi \rangle$ and the expression of the scalar product of spin coherent states given in²⁶ we get

$$\Phi = J \left[\phi - \varphi_{rf} - 2 \arctan \left[\cos \theta \tan \left(\frac{\phi - \varphi_{rf}}{2} \right) \right] \right]$$

In the classical method, the atom is initially prepared in $|J, J\rangle$, and the probability to find it in $|J, J\rangle$ at the end of the sequence is:

$$P(\varphi_{rf}) = |\langle \theta, \varphi_{rf} | \theta, \phi \rangle|^2 = \exp(-J \sin^2 \theta (\phi - \varphi_{rf})^2 / 2)$$

In order to measure Φ , we prepare, with a first microwave $\pi/2$ pulse, a quantum superposition of $|J, J\rangle$ and of the reference state $|R\rangle$. We then apply a second $\pi/2$ pulse resonant with the $|J, J\rangle \rightarrow |R\rangle$ transition after the rf pulses. The probability to find the atom in $|J, J\rangle$ is then given by:

$$P(\varphi_{rf}, \varphi_{mw}) = \frac{1}{4} + \frac{1}{4} P(\varphi_{rf}) + \frac{1}{2} \sqrt{P(\varphi_{rf})} \cos(\Phi - \varphi_{mw})$$

where φ_{mw} is the relative phase between the microwave pulses, and Φ implicitly depends on φ_{rf} .

The probability to find the atom in $|J, J\rangle$ therefore oscillates with Φ , with an amplitude proportional to $\sqrt{P(\varphi_{rf})}$. For small values of $\delta\omega$, this amplitude is maximum for $\varphi_{rf} \approx 0$ and then:

$$\Phi \approx \Phi_0 + J(1 - \cos \theta) \left(\frac{\partial \omega}{\partial F} \right) \delta F \tau$$

where Φ_0 is the phase accumulated for $\delta\omega = 0$.

Expression of the single atom sensitivity. The single-shot sensitivity is given by $\sigma_F^1 = (\partial P(\varphi_{rf}, \varphi_{mw}) / \partial F)^{-1} \sigma_P$ where σ_P is the dispersion of an atomic state detection. It can be rewritten as

$$\sigma_F^1 = \left| \frac{\partial P}{\partial \Phi} \right|^{-1} \left| \frac{\partial \Phi}{\partial F} \right|^{-1} \sigma_P \quad (5)$$

The optimum strategy to measure the electric field is to set the phase φ_{rf} that maximizes the contrast of the fringes, and to set φ_{mw} so that $P = 1/2$ (mid-fringe setting). Therefore $\sigma_P = \sqrt{P(1-P)} = 1/2$ is maximal.

In the ideal case, $\partial P / \partial \Phi = 1/2$, $\partial \Phi / \partial F = \tau J(1 - \cos \theta)(\partial \omega / \partial F)$, leading to a theoretical single-shot sensitivity:

$$\sigma_F^{1,th} = \frac{1}{\tau J(1 - \cos \theta)} \left| \frac{\partial \omega}{\partial F} \right|^{-1} \quad (6)$$

If $\theta = \pi$, the single shot sensitivity reaches the Heisenberg limit $\sigma_F^{1,HL} = (1/2\tau) |\partial \omega / \partial F|^{-1}$, which is $\sim \sqrt{J}$ smaller than the SQL (equation (1) in the main text). This improvement can be qualitatively analysed by considering Figs 2a and 3a. The classical phase $\delta\phi$ is the horizontal distance between the black and red Gaussian curves in Fig. 2a. The quantum phase $\delta\Phi$ is the horizontal distance between the black and red curves of Fig. 3a. It is J times larger than $\delta\phi$. However, the width of the Gaussian in Fig. 2a is \sqrt{J} times smaller than the period of the fringes in Fig. 3a. Therefore the gain in precision between the two methods is only \sqrt{J} (ref. 31).

Experimental single atom sensitivity. We calculate the experimental sensitivity corresponding to the interrogation time τ by considering only the differential phase $\delta\Phi_\tau$ accumulated between the two rf pulses, and writing $\partial \Phi / \partial F \approx \delta\Phi_\tau / \delta F$. We also take into account that $\partial P / \partial \Phi = C/2$ is reduced by the finite contrast C of the interference fringes. Finally, $\sigma_F^{1,exp} = 1/C \cdot \delta F / \delta\Phi_\tau$.

Physical interpretation of the spin J . To understand why the Rydberg atom behaves like a large spin when its evolution is restricted to the lower diagonal of the energy diagram of Fig. 1, it is helpful to use the Pauli replacement³² and express the restriction of the Stark Hamiltonian inside a given manifold in term of the Runge-Lenz vector \hat{A} (ref. 33). Then, in atomic units,

$$H_{\text{Stark}} = -\frac{3}{2} n F \hat{A}_z$$

Introducing the two operators $\hat{J}_1 = 1/2(\hat{L} + \hat{A})$ and $\hat{J}_2 = 1/2(\hat{L} - \hat{A})$, where \hat{L} is the orbital angular momentum operator, one can rewrite

$$H_{\text{Stark}} = -\frac{3}{2} n F (\hat{J}_{1,z} - \hat{J}_{2,z})$$

The operators \hat{J}_1 and \hat{J}_2 have equal magnitudes ($J_1 = J_2 = J = (n-1)/2$) and satisfy the commutation relation of two independent three-dimensional angular momenta³⁴. Therefore, the eigenstates of the Stark Hamiltonian are the states $|m_1, m_2\rangle = |J_1, M_1\rangle |J_2, M_2\rangle$, product states of the eigenstates associated with $\hat{J}_{1,z}$ and $\hat{J}_{2,z}$. Their energy is given by $E(m_1, m_2) = 3/2n(m_2 - m_1)F$. In this representation

$$|J, M\rangle = |J_1, m_1 = J_1\rangle |J_2, m_2 = M\rangle \quad (7)$$

The subset of the states $|J, M\rangle$ correspond to the restriction of the Stark manifold to the subspace where m_1 is maximum. In this subspace, $|J, M\rangle$ can be identified with the eigenstate of $\hat{J}_{2,z}$ such that $m_2 = M$. This is why, when its evolution is restricted to the $|J, M\rangle$ states, the Rydberg atom behaves like a single J -spin.

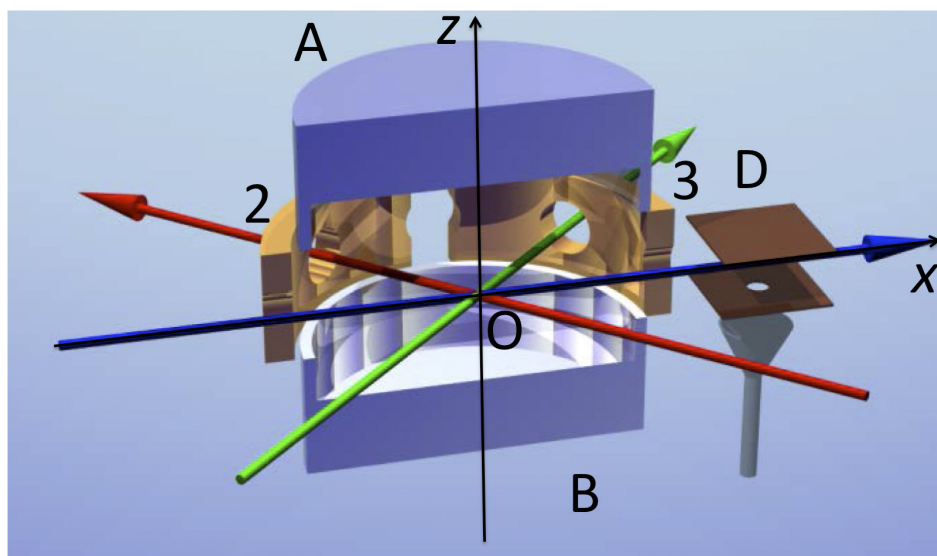
Finally, it is possible to use equation (7) to express the spin coherent state $|\theta, \phi\rangle$ of \mathbf{J} as a product of coherent states of \mathbf{J}_1 and \mathbf{J}_2

$$|\theta, \phi\rangle = |J\mathbf{u}_z\rangle |J\mathbf{u}_{\theta,\phi}\rangle$$

where \mathbf{u}_z and $\mathbf{u}_{\theta,\phi}$ are the unit vectors along the direction z and (θ, ϕ) . We use here the fact that the state $|J_1, m_1 = J_1\rangle$ is the spin coherent state along the z direction.

The spin coherent state $|\theta, \phi\rangle$ is thus an elliptic eigenstate of the hydrogen atom³³. The corresponding electronic wavefunction is localized along the Kepler ellipse defined by its angular momentum $\mathbf{L} = \langle \mathbf{J}_1 \rangle + \langle \mathbf{J}_2 \rangle = J\mathbf{u}_z + J\mathbf{u}_{\theta,\phi}$ and Runge-Lenz vector $\mathbf{A} = \langle \mathbf{J}_1 \rangle - \langle \mathbf{J}_2 \rangle = J\mathbf{u}_z - J\mathbf{u}_{\theta,\phi}$ (ref. 35). Starting from an horizontal circular orbit for $\theta = 0$, the classical trajectory becomes more and more elliptic, and more and more aligned with the z axis, as θ increases.

31. Bollinger, J. J., Itano, W. M., Wineland, D. J. & Heinzen, D. J. Optimal frequency measurements with maximally correlated state. *Phys. Rev. A* **54**, R4649–R4652 (1996).
32. Pauli, W. Jr Über das Wasserstoffspektrum vom Standpunkt der neuen Quantenmechanik. *Z. Phys.* **36**, 336–363 (1926).
33. Gay, J. C., Delande, D. & Bommier, A. Atomic quantum states with localization on classical elliptical orbits. *Phys. Rev. A* **39**, 6587–6590 (1989).
34. Bellomo, P., Stroud, C. R., Farrelly, D. & Uzer, T. Quantum-classical correspondence in the hydrogen atom in weak external fields. *Phys. Rev. A* **58**, 3896–3913 (1998).
35. Bellomo, P. & Stroud, C. R. Jr Classical evolution of quantum elliptic states. *Phys. Rev. A* **59**, 2139–2145 (1999).



Extended Data Figure 1 | Schematic of the experiment. The atoms are produced by excitation of a thermal rubidium beam (blue arrow) propagating along axis $O-x$. Two horizontal electrodes A and B (represented here as cut by a vertical plane) produce the directing electric field (F) along $O-z$. The gap between A and B is surrounded by four independent electrodes (1, 2, 3 and 4), on which we apply radio-frequency signals to produce σ_+ fields with tunable phase and amplitude. Electrodes

1 and 4, not represented, are the mirror images of electrodes 2 and 3 (in yellow) with respect to the $x-O-z$ plane. The laser excitation to the Rydberg states is performed using three laser beams that intersect in the centre O of the cavity. The 780 nm and 776 nm laser beams are collinear (red), the 1,259 nm laser is sent perpendicular to the other beams (green). Once the atoms have left the electrode structure, they enter the field-ionization detector D.

Transport evidence for Fermi-arc-mediated chirality transfer in the Dirac semimetal Cd_3As_2

Philip J. W. Moll^{1,2}, Nityan L. Nair¹, Toni Helm^{1,2}, Andrew C. Potter¹, Itamar Kimchi¹, Ashvin Vishwanath¹ & James G. Analytis^{1,3}

The dispersion of charge carriers in a metal is distinctly different from that of free electrons owing to their interactions with the crystal lattice. These interactions may lead to quasiparticles mimicking the massless relativistic dynamics of high-energy particle physics^{1–3}, and they can twist the quantum phase of electrons into topologically non-trivial knots—producing protected surface states with anomalous electromagnetic properties^{4–9}. These effects intertwine in materials known as Weyl semimetals, and in their crystal-symmetry-protected analogues, Dirac semimetals¹⁰. The latter show a linear electronic dispersion in three dimensions described by two copies of the Weyl equation (a theoretical description of massless relativistic fermions). At the surface of a crystal, the broken translational symmetry creates topological surface states, so-called Fermi arcs¹¹, which have no counterparts in high-energy physics or conventional condensed matter systems. Here we present Shubnikov–de Haas oscillations in focused-ion-beam-prepared microstructures of Cd_3As_2 that are consistent with the theoretically predicted ‘Weyl orbits’, a kind of cyclotron motion that weaves together Fermi-arc and chiral bulk states¹². In contrast to conventional cyclotron orbits, this motion is driven by the transfer of chirality from one Weyl node to another, rather than momentum transfer of the Lorentz force. Our observations provide evidence for direct access to the topological properties of charge in a transport experiment, a first step towards their potential application.

The bulk electrons in topological semimetals are described by an ultra-relativistic dispersion relation, $E(\mathbf{k}) = \pm \hbar v_F \boldsymbol{\sigma} \cdot \mathbf{k}$, that resembles the Weyl equation for massless spin-1/2 particles. Here E is electron energy, v_F the Fermi velocity, and $\boldsymbol{\sigma}$ is a pseudo-spin-1/2 degree of freedom that is energetically locked parallel or anti-parallel to the momentum, \mathbf{k} , of the electron, giving electrons definite chirality $\mathbf{k} \sim \pm \boldsymbol{\sigma}$. Applying electromagnetic fields to Weyl or Dirac semimetals induces a pumping of electric charge between Weyl nodes with opposite chirality, a phenomena known in high-energy physics as the chiral anomaly^{13–15}. At the surface of these materials, this anomalous chirality transfer is facilitated by topologically protected surface arcs, the so-called Fermi-arc surface states, which act as a pipeline connecting opposite chirality Weyl points^{11,16}. Recently, Na_3Bi (ref. 17) and Cd_3As_2 (ref. 18) have been predicted to be three-dimensional bulk Dirac semimetals. This has sparked substantial research interest and the linear dispersion in these materials has been confirmed by ARPES^{19–24} (angle-resolved photoemission spectroscopy) and STM²⁵ (scanning tunnelling microscopy) experiments. A number of unusual material properties such as strong linear magnetoresistance^{19,26} and high mobilities²⁷ were identified that are potentially linked to the relativistic nature of the Dirac quasiparticles. Yet the prospect of studying ultra-relativistic particles and their accompanying topological surface states, as well as potential applications exploiting their unusual behaviour, naturally requires a more direct measurement capable of revealing both the relativistic dynamics and topological surface states.

The main aim of this study is to present such evidence in four-terminal transport measurements, showing the possibility of detecting and manipulating chiral states in Dirac semimetals.

One intriguing fingerprint of Weyl quasiparticles in the electronic transport properties in strong magnetic fields has been recently predicted¹². The ‘Weyl orbit’ (Fig. 1) weaves together the chiral states in the bulk with the topological Fermi-arc states on opposite surfaces into a closed orbit. Its quantization produces a distinctive contribution to the quantum oscillation spectrum that provides an observable signature of the chiral and topological character of these materials. This closed orbit is strikingly different from typical electrons orbiting around a Fermi surface in a metal, as the quasiparticle experiences zero Lorentz force on the chiral path segments traversing the bulk.

The main result of this study is an additional quantum oscillation frequency observed in microstructures smaller than the mean free path that exhibits characteristics of both surface-like and bulk-like states, as naturally expected for Weyl orbits. The microstructures were prepared from Cd_3As_2 single crystals by focussed ion beam (FIB) etching (Fig. 1, Methods). Down to the smallest thickness of $L = 150$ nm, the magnetoresistance at temperatures below 100 K shows pronounced Shubnikov–de Haas oscillations signalling the low effective mass of the charge carriers and the high crystal quality of the devices (Fig. 1b). Studies of quantum oscillations in bulk crystals have reported one single frequency^{19,26}, arising from an essentially spherical three-dimensional (3D) Fermi surface in agreement with ARPES and STM experiments. This single bulk frequency (F_B) is also consistently observed in all of the studied parent bulk crystals as well as all microstructures. However, when the field is applied perpendicular to the [010] surface (denoted 0° , see below), a second frequency $F_S = 61.5$ T appears which is distinct from the higher harmonics of the bulk. We find an effective mass of $0.044m_e$ (where m_e is the electron’s mass) for the bulk, similar to previously reported measurements on bulk crystals²⁶, and a similar mass of $0.050m_e$ in the additional orbit F_S (see Methods).

The value of the observed surface frequency F_S is compatible with the prediction for Weyl orbits. Their trajectory combines segments of both chiral bulk and Fermi-arc surface states, and F_S can be estimated from the time spent in each of them¹²:

$$F_S = E_F k_0 / (e\pi v_F) \approx 56 \text{ T}$$

The length of the surface Fermi arc in reciprocal space may be approximated as $k_0 \approx 0.8 \text{ nm}^{-1}$ using the k -space separation of the Dirac points determined in ref. 19. The bulk quantum oscillation frequency and effective mass provide direct access to the Fermi energy $E_F = 192 \text{ meV}$ and the Fermi velocity $v_F = 8.8 \times 10^5 \text{ m s}^{-1}$. The resulting estimate of $F_S \approx 56 \text{ T}$ is in good quantitative agreement with the measured $F_S = 61.5 \text{ T}$.

The dependence of the quantum oscillation frequencies on the angle between the magnetic field and the surface normal shows a

¹Department of Physics, University of California, Berkeley, California 94720, USA. ²Max Planck Institute for Chemical Physics of Solids, Noethnitzer Strasse 40, D-01187 Dresden, Germany.

³Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.

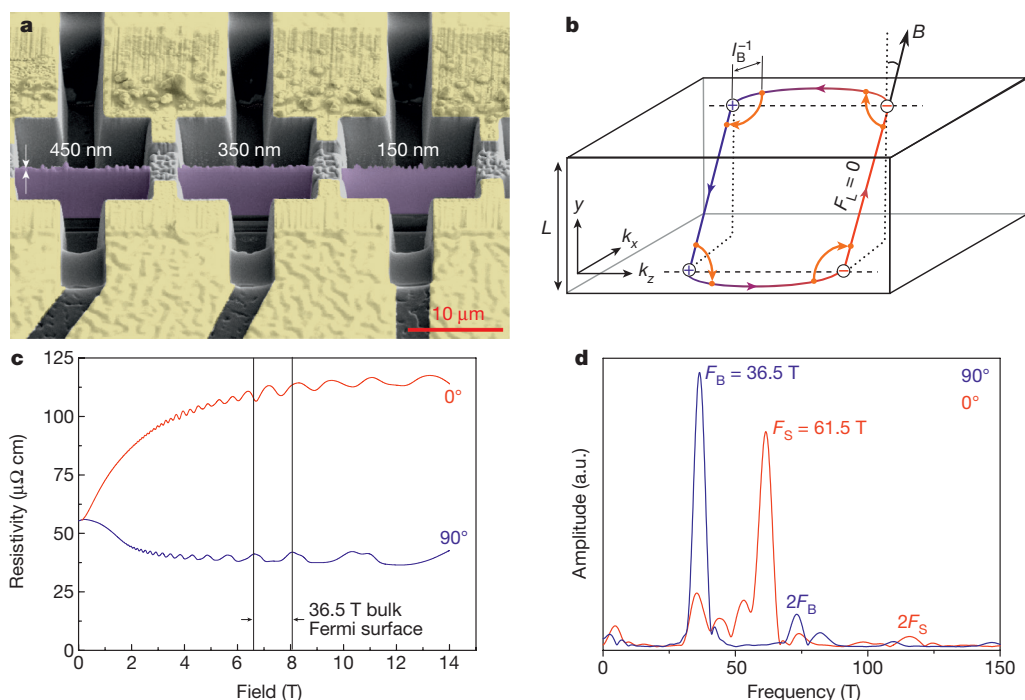


Figure 1 | Surface oscillations in microstructures. **a**, SEM micrograph of a typical sample prepared by FIB cutting. More than 10 similar devices have been fabricated and the results were highly reproducible among all of them. The active devices are $4\mu\text{m} \times 10\mu\text{m}$ free-standing sheets (purple) of varying thickness connected to contact pads (yellow). The crystallographic direction perpendicular to the polished surface is [010] and the direction parallel to it is [100], which define the plane of rotation. This entire sample consists of one contiguous slice of a Cd_3As_2 crystal, ensuring that all devices are made from the exact same starting material and have the same orientation during the experiments. **b**, Sketch of the Weyl orbit in a thin slab of thickness L in a magnetic field B . The orbit involves both the Fermi-arc surface states connecting the Weyl nodes of opposite chirality, and the bulk states of fixed chirality (blue and red). In strong magnetic

fields, quasiparticles may tunnel through the energy barrier separating the bulk states from the surface over a distance associated with the magnetic length l_B (orange arrows). Note that as the Weyl orbit mixes processes in real and reciprocal space, the (x, y) coordinates of this sketch are in reciprocal space and the z coordinate is in real space. **c**, Magnetoresistance and **d**, its Fourier transform measured on the thinnest device (150 nm) at 2 K for fields parallel (90°) and perpendicular (0°) to the surface. The main finding of this study is directly evident in the raw data: while parallel fields lead to a single frequency (including spin-splitting at higher fields), an additional higher frequency component F_S associated with the surface oscillations appears for perpendicular fields. The frequency difference between the two directions becomes evident in the raw data when comparing the resistivity peaks directly, as indicated by the line markers.

clear signature of a surface state, as shown in Fig. 2, where 0° denotes fields perpendicular to the surface and 90° parallel to it. The angular dependencies of both frequencies are strikingly different: while the low frequency F_B (blue) remains essentially constant and is observed at all angles, the frequency F_S (red) strongly increases as the field is tilted away from the surface. $F_S(\theta)$ is well described by a $\cos(\theta)^{-1}$ dependence, indicating that the field component perpendicular to the surface is relevant for the orbit. This angle dependence is a hallmark of two-dimensional (2D) Fermi surfaces. The bulk band structure however does not support a 2D Fermi surface and it is not observed in our bulk crystals.

While the angle dependence clearly suggests a surface character of the quantum path associated with the frequency F_S , the quantum oscillations also show pronounced bulk-like characteristics which are unexpected for orbits simply consisting of surface states. First of all, the additional quantum oscillation is only observed in samples where the bulk mean free path is longer than or comparable to the sample thickness. The low-field transverse magnetoresistance for in-plane fields quantitatively confirms our devices to be in this limit. The resistance maximum at small fields is a hallmark signature of quasi-ballistic transport in clean metals known as the Knudsen effect (Fig. 3a, Methods).

The amplitudes of the quantum oscillations are found to be strongly thickness dependent (Fig. 3b). No trace of the surface frequency F_S has been observed in devices thicker than $3\mu\text{m}$, in agreement with its absence in our bulk quantum oscillation measurements. As the sample thickness is reduced, its relative weight compared to the bulk

frequency strongly increases, and devices thinner than 500 nm are dominated by the surface oscillation. The increase in the various studied samples follows an exponential behaviour, with an exponent of $d = 675\text{ nm}$. This value should be compared to the bulk mean free path estimated from transport as $l = v_F\tau = v_F m^* / (ne^2 \rho_0) = 1.0\mu\text{m}$, where $\rho_0 = 55\mu\Omega\text{ cm}$ is the zero-field resistivity in our crystals at 2 K, $n = 2k_F^3 / (3\pi^2) = 2.5 \times 10^{18}\text{ cm}^{-3}$ the bulk carrier density estimated for the twofold degenerate spherical Fermi surface, τ the scattering time, m^* the effective mass of the charge carriers, and $k_F = m^* v_F / \hbar = 3.3 \times 10^8\text{ m}^{-1}$ the Fermi momentum extracted from the Shubnikov–de Haas oscillations. This estimate of the mean free path is in good quantitative agreement with the observation of the Knudsen flow maximum. Ordinarily, any scattering process leads to quantum decoherence, in which case the quantum oscillations would be expected to decay exponentially in the device thickness L over the quantum mean free path l_Q which is generally much shorter than the mean free path measured in transport. However, for large fields the bulk chiral Landau level is expected to exhibit extra resilience to such dephasing and allow for quantum oscillations for thicker samples. As the closed Weyl orbit contains two path segments traversing the bulk, each at opposite chirality, the effective bulk path length is twice the device thickness L . The agreement between $2d = 1.35\mu\text{m}$ (as $d = 0.675\mu\text{m}$) and $l = 1.0\mu\text{m}$ suggests that the relevant thickness scale for the appearance of the additional quantum oscillations is comparable to the bulk transport mean free path.

An additional clue to the origin of the Shubnikov–de Haas oscillations comes from their phase as a function of field. We observe

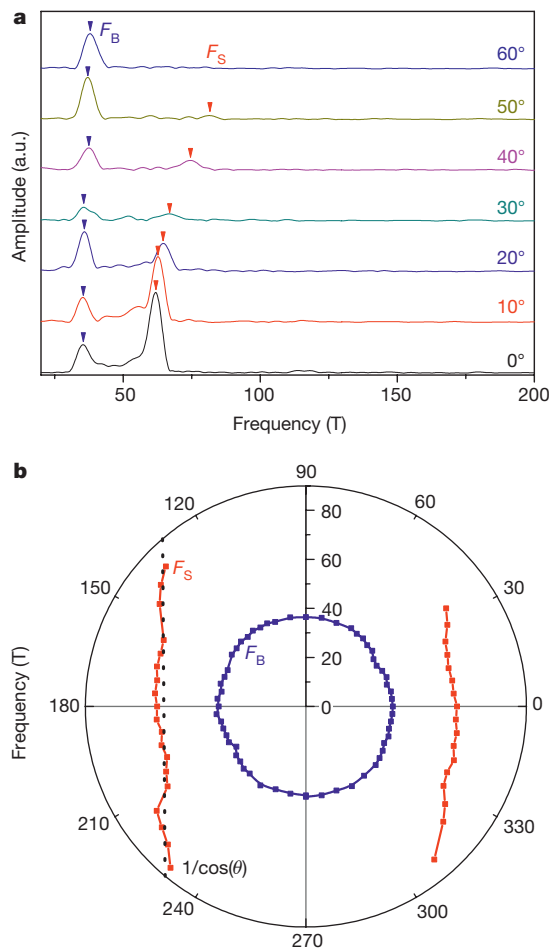


Figure 2 | Angle-dependent oscillations. **a**, Angle dependence of the quantum oscillation spectrum measured in the 150 nm device. As the field is rotated away from the perpendicular configuration (0°), the surface frequency F_S increases and shrinks in amplitude, while the bulk frequency is unaffected. **b**, Polar plot of the angle dependence for the bulk F_B and the surface frequency F_S . F_B is almost isotropic while F_S increases as the field is rotated towards a parallel configuration. It follows a surface-like $\cos(\theta)^{-1}$ dependence represented by a straight, off-centre line in a polar plot.

deviations from ideal periodicity of the oscillations in inverse magnetic field, appearing as a continuously drifting phase as shown in Fig. 4. While similarly subtle deviations from periodicity can arise from g -factor band splitting in strongly spin–orbit coupled materials²⁸, the direction of the shift is opposite to those previously observed in experiments: the peak positions B_n associated with the n th Landau level are shifted towards higher fields compared to the purely periodic case, while spin-splitting should shift them towards lower fields. On the other hand, this direction and magnitude of shift of the Landau levels is expected for Weyl orbits. Non-adiabatic corrections are expected to appear because of field-induced tunnelling between Fermi-arc states and bulk states, occurring as the orbiting quasiparticles approach the Weyl node. A single Weyl orbit will encounter four such tunnelling processes, leading to a phase deviation that is in quantitative agreement with what we observe (Fig. 4, Methods).

The discussion so far has considered the Dirac material Cd_3As_2 as two independent Weyl subsystems overlapping in k -space. However, magnetic fields can break the crystal symmetry protecting the superimposed Weyl nodes of opposite chirality and thus introduce a gap. In sufficiently strong fields, the particle approaching the node may thus tunnel into the oppositely dispersing Fermi arc on the same surface instead of into the bulk, thus forming closed orbits purely from Fermi-arc states, similar to the surface states of topological insulators. While

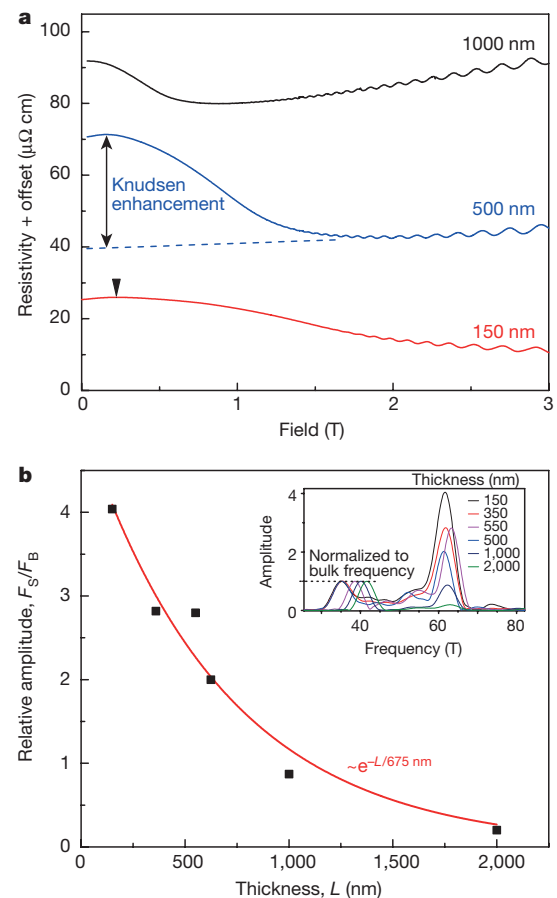


Figure 3 | Thickness dependence. **a**, A magnetoresistance maximum was observed in all studied samples for fields applied parallel to the surface (90°). This peak arises from the semi-classical Knudsen effect observed in quasi-ballistic transport when the cyclotron radius becomes comparable to the sample dimensions. **b**, Relative amplitude of the surface oscillations compared to the bulk oscillations, for fields perpendicular to the surface at 2 K. The surface oscillations are unobservable for devices thicker than $3\ \mu\text{m}$, and their relative weight grows with thinning of the samples. Inset, raw frequency spectra for devices of different thickness.

our experiments cannot rule out localized topological surface states as the origin of the oscillations, the thickness dependence of the surface state amplitudes and the non-adiabatic corrections are more suggestive of the Weyl orbit in the present field range. In both cases, however, the Fermi arcs take part in the quantum orbit and our results thus confirm the detection of topological surface currents in the microstructures.

However, quantum oscillations may also arise from trivial surface states or from the defect layer introduced by the fabrication technique, without the involvement of topological states. Thus it is essential to find ways to determine if topological states participate in the observed orbit, setting the involved orbit apart from topologically trivial states. The existence of a ‘saturation field’ above which oscillations were expected to cease was predicted to provide such a test¹². However, a re-examination of the equation for quantum oscillations, reproduced in equation (1) below, reveals that the oscillations associated with the Weyl orbit actually persist up a much larger field of order $F_S = 56\ \text{T}$ in the present instance (Methods), and the thickness dependence of the saturation field is highly complex. Consequently, the Weyl surface oscillations are expected to persist over the full range of fields explored in this experiment, in agreement with our observations. Additional evidence for the topological nature of the surface oscillations comes from their remarkable resilience against surface disorder. We have purposely introduced strong surface damage by almost normal incidence

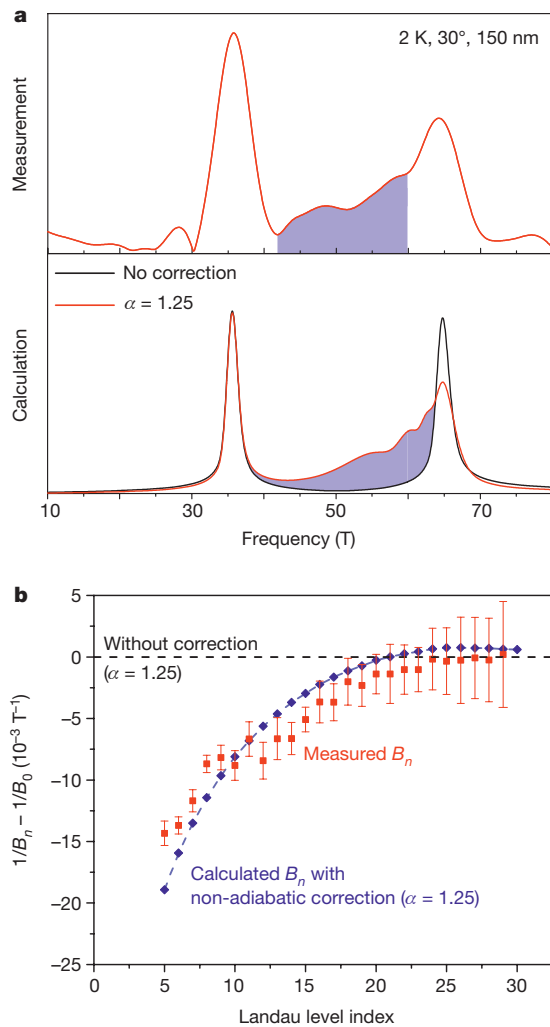


Figure 4 | Non-adiabatic corrections. **a**, Comparison of a measured oscillation spectrum (top panel) to simulations (bottom panel), with and without the non-adiabatic correction term. Without non-adiabatic corrections (black trace), the expected spectra are well-defined peaks of similar broadening. The non-adiabatic corrections (red trace) lead to an asymmetric reduction of the effective amplitude in high magnetic field, and shift spectral weight from the main surface peak to lower frequencies (shaded area). **b**, Red squares, difference between the observed resistance maxima corresponding to the n th Landau level, $1/B_n$, and the positions $1/B_0$ extrapolated from low fields using the usual $1/B$ periodicity. The difference can be well explained by including the non-adiabatic correction term using $\alpha = 1.25$ (blue diamonds)¹². A signal periodic in $1/B$ in the absence of non-adiabatic corrections would fall onto the dashed line. The error bars denote the experimental uncertainty of the peak positions.

FIB irradiation. Remarkably, the surface state oscillations grow upon increasing disorder, in strong contrast to the usual disorder dampening of conventional surface oscillations. The counter-intuitive increase of amplitude with increasing surface damage, however, would be a natural consequence of the protection of topological surface states (Methods).

Another striking piece of evidence for the non-trivial nature of the quantum orbit comes from a distinguishing feature of the Fermi-arc orbit: The position of the n th resistance minimum, B_n , depends on the thickness of the sample, L , given by:

$$B_n^{-1} = ek_0^{-1}(n\pi v_F/E_F - L) \quad (1)$$

To test this concept, we fabricate a sample geometry that is by design sensitive to such a thickness dependence of the phase. In Fig. 5 we show a sample structured with a triangular cross-section as well as

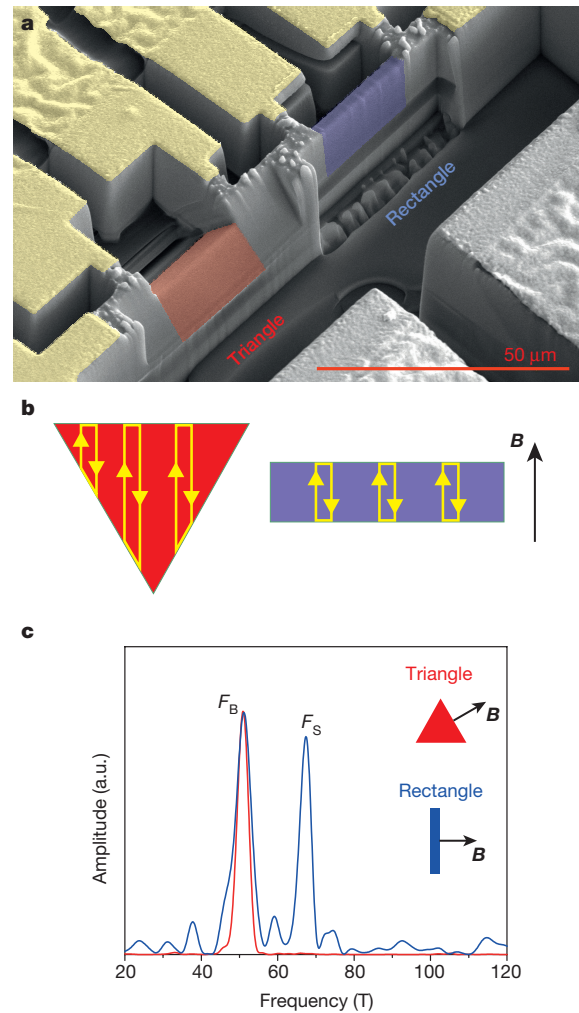


Figure 5 | Absence of surface oscillations in triangular samples. **a**, SEM image of triangular and rectangular devices. The rectangular sample is $L = 0.8 \mu\text{m}$ wide, $3.2 \mu\text{m}$ tall and $5 \mu\text{m}$ long. The other device features an equilateral triangular cross-section with a base of $a = 2.7 \mu\text{m}$. Both devices have a similar cross-sectional area (rectangle, $0.8 \times 3.2 = 2.6 \mu\text{m}^2$; triangle, $a = 3.1 \mu\text{m}^2$) and circumference of the cross-section (rectangle, $2 \times (3.2 + 0.8) = 8.0 \mu\text{m}$; triangle, $8.1 \mu\text{m}$). The crystallographic direction perpendicular to the surface of the rectangular device is $[102]$ and $[010]$ parallel to the surface, which defines the plane of rotation. **b**, Sketch of the Weyl orbits for rectangular and triangular cross-sections. **c**, Frequency spectrum of the triangular and rectangular samples, for field orientations perpendicular to each of the surfaces (0° for the rectangle and 60° for the triangle).

a rectangular one as a reference. The quantum oscillations for fields perpendicular to the surface for each device (0° for the rectangle, 60° for the equilateral triangle) are strikingly different. While the rectangular device clearly shows the presence of a surface state, the triangular one shows only the bulk frequency without any sign of the surface frequency. Crucially, all surfaces were fabricated under exactly the same conditions from the same piece of crystal. The sample was tilted with respect to the ion beam to ensure a grazing incidence condition for every surface, both on the rectangular and the triangular device. Both devices have comparable cross-section and surface area by design, so that a trivial surface state acting as a parallel conductance channel should lead to observable surface quantum oscillations in both of them. The absence of the surface state in the triangle is unexpected for oscillations arising from a trivial surface state, yet a natural consequence of Weyl orbits. Unlike conventional surface states that are confined to a single surface, the Weyl orbit cannot be observed in triangular geometries: all paths are of different length, each contributing to the field induced density-of-state modulation at a different magnetic field $1/B_n$.

(see equation (1), Fig. 5). This results in destructive interference due to a sum of oscillations with random phases, rendering the quantum oscillations unobservable in an experiment. In contrast, all quantum paths involve bulk path segments of the same length L in a rectangular geometry, and thus all contribute to a density-of-state modulation at the same field. This is direct evidence that the orbit associated with the frequency F_S is sensitive to the shape and size of the bulk underneath the surface, in contrast to those arising from trivial surface states.

The ensemble of presented results highlights an essential aspect of the additional quantum oscillations that appear when Dirac fermions are confined within microstructures smaller than the mean free path: they share characteristic features of both surface-like and bulk-like oscillations. Such hybrid characteristics arise naturally from the idea of the mixing of chiral bulk-states and Fermi-arc surface-states into a coherent orbit. Nevertheless, there remain questions that challenge our current understanding of topological matter on microscopic length scales and in strong magnetic fields. For example, understanding the exact amplitude of the Weyl oscillations may require the extension of the present theory into the quantum limit, where the number n of occupied Landau levels cannot be treated as being large. Furthermore, our understanding of the mixing of chiral states in strong magnetic fields and their influence on the Weyl orbit needs to be improved and experimentally investigated in the future. By using FIB structuring, we have demonstrated an experimentally simple route towards studying strongly confined topological matter, which will both increase our understanding of the transport characteristics of Fermi-arc states and provide a path to investigate the potential of these materials in future electronic applications.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 20 November 2015; accepted 13 April 2016.

Published online 4 July 2016.

- Katsnelson, M. I., Novoselov, K. S. & Geim, A. K. Chiral tunnelling and the Klein paradox in graphene. *Nat. Phys.* **2**, 620–625 (2006).
- Novoselov, K. S. *et al.* Two-dimensional gas of massless Dirac fermions in graphene. *Nature* **438**, 197–200 (2005).
- Vafeek, O. & Vishwanath, A. Dirac fermions in solids — from high T_c cuprates and graphene to topological insulators and Weyl semimetals. *Annu. Rev. Condens. Matter Phys.* **5**, 83–112 (2014).
- Fu, L., Kane, C. L. & Mele, E. J. Topological insulators in three dimensions. *Phys. Rev. Lett.* **98**, 106803 (2007).
- Fu, L. & Kane, C. Superconducting proximity effect and Majorana fermions at the surface of a topological insulator. *Phys. Rev. Lett.* **100**, 096407 (2008).
- Roushan, P. *et al.* Topological surface states protected from backscattering by chiral spin texture. *Nature* **460**, 1106–1109 (2009).
- Zhang, T. *et al.* Experimental demonstration of topological surface states protected by time-reversal symmetry. *Phys. Rev. Lett.* **103**, 266803 (2009).
- Pan, Z.-H. *et al.* Electronic structure of the topological insulator Bi_2Se_3 using angle-resolved photoemission spectroscopy: evidence for a nearly full surface spin polarization. *Phys. Rev. Lett.* **106**, 257004 (2011).
- Park, S. R. *et al.* Chiral orbital-angular momentum in the surface states of Bi_2Se_3 . *Phys. Rev. Lett.* **108**, 046805 (2012).
- Young, S. M. *et al.* Dirac semimetal in three dimensions. *Phys. Rev. Lett.* **108**, 140405 (2012).
- Wan, X., Turner, A. M., Vishwanath, A. & Savrasov, S. Y. Topological semimetal and Fermi-arc surface states in the electronic structure of pyrochlore iridates. *Phys. Rev. B* **83**, 205101 (2011).
- Potter, A. C., Kimchi, I. & Vishwanath, A. Quantum oscillations from surface Fermi arcs in Weyl and Dirac semimetals. *Nat. Commun.* **5**, 5161 (2014).
- Adler, S. L. Axial-vector vertex in spinor electrodynamics. *Phys. Rev.* **177**, 2426–2438 (1969).
- Bell, J. S. & Jackiw, R. A PCAC puzzle: $\pi^0 \rightarrow \gamma\gamma$ in the σ -model. *Nuovo Cim. A* **60**, 47–61 (1969).
- Nielsen, H. B. & Ninomiya, M. The Adler-Bell-Jackiw anomaly and Weyl fermions in a crystal. *Phys. Lett. B* **130**, 389–396 (1983).
- Haldane, F. D. M. Attachment of surface “Fermi arcs” to the bulk Fermi surface: “Fermi-level plumbing” in topological metals. Preprint at <http://arXiv.org/abs/1401.0529v1> (2014).
- Wang, Z. *et al.* Dirac semimetal and topological phase transitions in A_3Bi ($\text{A}=\text{Na}, \text{K}, \text{Rb}$). *Phys. Rev. B* **85**, 195320 (2012).
- Wang, Z., Weng, H., Wu, Q., Dai, X. & Fang, Z. Three-dimensional Dirac semimetal and quantum transport in Cd_3As_2 . *Phys. Rev. B* **88**, 125427 (2013).
- Liang, T. *et al.* Ultrahigh mobility and giant magnetoresistance in the Dirac semimetal Cd_3As_2 . *Nat. Mater.* **4**, 3–7 (2014).
- Neupane, M. *et al.* Observation of a three-dimensional topological Dirac semimetal phase in high-mobility Cd_3As_2 . *Nat. Commun.* **5**, 3786 (2014).
- Liu, Z. K. *et al.* A stable three-dimensional topological Dirac semimetal Cd_3As_2 . *Nat. Mater.* **13**, 677–681 (2014).
- Yi, H. *et al.* Evidence of topological surface state in three-dimensional Dirac semimetal Cd_3As_2 . *Sci. Rep.* **4**, 6106 (2014).
- Liu, Z. K. *et al.* Topological Dirac semimetal, Na_3Bi . *Science* **343**, 864–867 (2014).
- Borisenko, S. *et al.* Experimental realization of a three-dimensional Dirac semimetal. *Phys. Rev. Lett.* **113**, 027603 (2014).
- Jeon, S. *et al.* Landau quantization and quasiparticle interference in the three-dimensional Dirac semimetal Cd_3As_2 . *Nat. Mater.* **13**, 851–856 (2014).
- He, L. P. *et al.* Quantum transport evidence for the three-dimensional Dirac semimetal phase in Cd_3As_2 . *Phys. Rev. Lett.* **113**, 246402 (2014).
- Rosenberg, A. J. & Harman, T. C. Cd_3As_2 — a noncubic semiconductor with unusually high electron mobility. *J. Appl. Phys.* **30**, 1621–1622 (1959).
- Analytis, J. G. *et al.* Two-dimensional surface state in the quantum limit of a topological insulator. *Nat. Phys.* **6**, 960–964 (2010).

Acknowledgements The FIB work was supported by the SCOPE-M centre for electron microscopy at ETH Zurich. We thank P. Gasser, J. Reuteler and B. Batlogg for FIB support, and M. Bachmann for performing magnetoresistance measurements. A.C.P. was supported by the Gordon and Betty Moore Foundation’s EPIQS Initiative through grant GBMF4307. We also thank S. Teat and K. Gagnon for their help in conducting X-ray diffraction measurements at the Advanced Light Source (ALS) beam line 11.3.1, and N. Tamura for micro-diffraction on beam line 12.3.2. N.T. and the ALS are supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division, of the US Department of Energy under contract no. DE-AC02-05CH11231 at Lawrence Berkeley National Laboratory and the University of California, Berkeley. Transport experiments, material synthesis and FIB microstructuring were supported by the Gordon and Betty Moore Foundation’s EPIQS Initiative through grant GBMF4374. Single-crystal X-ray refinements (T.H.) and theoretical support (A.V. and I.K.) were funded by the Quantum Materials FWP, US Department of Energy, Office of Basic Energy Sciences, Materials Sciences and Engineering Division, under contract no. DE-AC02-05CH11231. Work at the Molecular Foundry was supported by the Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under contract no. DE-AC02-05CH11231.

Author Contributions P.J.W.M. microstructured the crystals, and performed the measurements and data analysis. N.L.N. synthesized and characterized the single crystals. T.H. analysed the crystal structure. I.K., A.C.P. and A.V. contributed the theoretical treatment. P.J.W.M. and J.G.A. designed the experiment. All authors were involved in writing the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.J.W.M. (philip.moll@cpfs.mpg.de) or J.G.A. (analytis@berkeley.edu).

METHODS

FIB sample preparation. The focused ion beam (FIB) has proven to be a powerful tool to fabricate crystalline microstructures of high quality (see, for example, refs 29–33) to study meso- and microscopic transport phenomena such as the Weyl-orbit quantum oscillations presented in this study. In this process, the microstructures are carved out of macroscopic crystals using the FIB. Starting from millimetre-sized flux-grown single crystals presented in Extended Data Fig. 1, we use a 30-kV Ga^{2+} ion beam to cut the crystal. Depending on the ion flux required, the irradiation spot can be as small as a few nm, and thus structures can be fabricated with high precision. The beam is guided across a negative image of the desired structure, which defines the parts of the crystal to be removed. Electrical contacts to the device can also be made *in situ* in the FIB. A platinum precursor gas can be introduced into the chamber, and upon interaction of the ion beam with the gas adsorbed to the surface a conductive, Pt-rich film can be grown connecting the microstructure electrically to external leads.

The fabrication occurs in two steps: in the first step, a cross-sectional slice is cut from the crystal. Typical dimensions of this slice are $100\text{ }\mu\text{m} \times 15\text{ }\mu\text{m} \times 3\text{--}4\text{ }\mu\text{m}$. Cd_3As_2 is found to be sputtered very effectively by the ion beam and even large volumes can be sputter etched with low ion currents at a reasonable rate. This renders Cd_3As_2 an ideal material for FIB microstructuring and the low currents contribute to the high crystal quality of the fabricated devices. For the coarse patterning in this step, currents of 2.5 nA were used. This slice is then removed *ex situ* under an optical microscope and transferred to a Si/SiO_2 substrate with predefined gold leads connecting the centre to bonding pads. The slice is glued in the centre using a small drop of epoxy and cured for 1 h at 100°C . Good results were obtained with both Araldite Rapid and Stycast 1266 epoxies. Cd_3As_2 is found to react adversely with the platinum precursor gas leading to an insulating interface with non-ohmic behaviour between the platinum contacts and the crystal slice. To avoid this barrier formation, the slice was sputter-coated in 100-nm gold, which was later removed by FIB etching from the top-side of the studied microdevices. A more detailed description of the process was previously published in ref. 34. In the second step, the mounted slice is again introduced into the FIB for final patterning into the thin bar shape shown in the main text. Again the high sputter rate allowed us to use currents as low as 40 pA for the structuring process. First, the coarse outline of the final structure was patterned at 300 pA, and it was checked that no stray connection due to redeposited material remained. Then the final thickness was adjusted by 40 pA milling. The scan strategy for the final approach was a line-by-line, single pass mill from both sides of the device to ensure symmetric, highly polished surfaces. The strong bulk-quantum oscillations indicate the high crystallinity of our microstructures, which was additionally confirmed via X-ray microdiffraction at the Advanced Light Source (Lawrence Berkeley National Laboratory).

Nature of the surface layer. FIB microstructuring is a very gentle process inducing little bulk damage as evidenced by the strong bulk quantum oscillations present in our Cd_3As_2 devices. On the other hand, the impacting ions are known to damage the crystal surface and create an amorphous surface layer. Therefore particular care must be taken to investigate if a highly conductive layer is formed that exhibits conventional surface Shubnikov–de Haas oscillations. As discussed in the main text, the experimental evidence points against a trivial conductive surface layer as the source of the observed surface-like oscillations: (1) these heavily disordered damage layers are unlikely to show quantum oscillations due to their short mean free path, (2) it would not show adiabatic corrections to the quantum oscillation frequency and (3) the triangular device shows no sign of surface oscillations despite its comparable surface area and crystal orientation. Yet it remains essential to understand the surface layer, and therefore we analyse its expected thickness and nature by Monte Carlo simulations of ion interaction with Cd_3As_2 , which provide an estimate of the expected surface layer thickness around 20 nm. The interaction of ions with matter can be well captured using Monte-Carlo simulations. We have performed a damage analysis including a calculation of the full cascade damage using the SRIM package (Stopping Range of Ions in Matter)³⁵. Under realistic fabrication conditions of grazing incidence polishing (88° to surface normal), we found a longitudinal stopping range of 7.6 nm in Cd_3As_2 . This parameter describes the mean implantation depth of Ga, with a standard deviation of 5.4 nm. At grazing incidence, the Ga-ion backscattering probability is strongly enhanced and thus the Ga-implantation is suppressed compared to normal incidence. The main nature of the defects in the surface layers are cascade damages, resulting in a displacement of Cd and As atoms from their positions in the crystal lattice. The first 20 nm of material surface are thus expected to be amorphous, with an exponential reduction of defects and damage deeper in the bulk. This is in agreement with the observation of quantum oscillations corresponding to the bulk frequency in our samples.

In the following we investigate more directly if indeed topological surface states are involved in the surface-like orbit: one consequence of the topological

character of the Fermi-arc states involved in the Weyl orbit is their protection from surface disorder. This can be directly tested by purposely damaging the FIB-polished surfaces of Cd_3As_2 . By irradiating one side of a device at close-to-normal beam incidence, the surface layer can be severely damaged and disordered. Upon increasing surface damage, the surface state quantum oscillations of a trivial surface state are expected to be dampened due to the decreased quasiparticle lifetime following the Lifshitz–Kosevich mechanism³⁶. On the other hand, the topological surface state existing at the boundary between the periodic crystal bulk and the amorphous surface layer would remain unaffected, as it is simply pushed deeper into the bulk. To inflict maximal surface damage, exceptionally long beam dwell times at 30° angle between the beam and the surface normal were used to increase heating and cause deeper defect penetration (dwell time 1 ms, current 80 pA, acceleration voltage 30 kV). Under these conditions, an easily visible pattern of shallow dimples appears on the damaged surface (Extended Data Fig. 2b). The structure is of the same design as discussed in the main text with a thickness of 810 nm from the same Cd_3As_2 crystal, yielding a ratio of quantum oscillation amplitudes of bulk and surface states of 1.2 as expected at this thickness (Fig. 3). The same device was fabricated with polished surfaces and first well characterized in its pristine state. Then the same device was measured again after the irradiation, thus allowing direct investigation of the changes due to the irradiation.

The resistivity of the device increased by 55%, indicating the effective introduction of surface disorder due to the irradiation (Extended Data Fig. 2c). Remarkably, however, the amplitude of the Shubnikov–de Haas oscillations of both the surface and the bulk states have increased due to the irradiation (Extended Data Fig. 2d, e). This is in stark contrast to the expected strong reduction of amplitude due to disorder of a trivial surface state. However, this unusual result is a natural consequence of a topologically protected surface state. The current in such structures is expected to flow in three parallel paths: the crystal bulk, the topological surface state, and the disordered surface layer. By irradiating the surface at low incidence angle, the outermost damage layer becomes more disordered, as evidenced by the increase of the total device resistance. Therefore, the relative weight of the current flowing through the topological states and the bulk increases, leading to the unusual phenomenon of increasing surface state quantum oscillations upon increasing surface damage. This is further supported by the almost unchanged surface-to-bulk amplitude ratio due to irradiation (Extended Data Fig. 2e).

The robustness of the observed surface state quantum oscillations to artificial extensive surface damage strongly suggests a mechanism at play that protects the surface state from excess scattering, which would arise naturally from the topological properties of the Fermi-arc states. This effect of topological protection from non-magnetic scattering has been shown in Cd_3As_2 (refs 20, 22) as well as in topological insulators^{6,7,37}. Clearly, such robust quantum coherent signatures are contrary to the expectations for trivial surface states.

Quantum oscillation analysis. Magnetic fields quantize the electronic degrees of freedom orthogonal to the field direction. The energy of the Landau levels grows with increasing field given by equation (2), and the Landau levels are successively emptied as they are pushed above the Fermi level. This leads to a well-known modulation of the density-of-states at the Fermi level (DOS), and conversely to an oscillatory behaviour of physical quantities depending on the DOS in high fields. A good overview of such quantum oscillatory effects is given in ref. 36. The effective mass m^* is obtained from the temperature dependent dampening of the quantum oscillations due to thermal broadening of the Landau levels. The thermal dampening is described by the Lifshitz–Kosevich formalism and it can be shown that the oscillation amplitude follows the functional form $A(T)/A_0 = X/\sinh(X)$ where $X = 2\pi^2 k_B T/(\hbar\omega_c)$. Here k_B denotes the Boltzmann constant and $\omega_c = eB/m^*$ is the cyclotron frequency. From fitting this temperature dependence to the measured amplitudes, as shown in Extended Data Fig. 3a, the effective mass is obtained as a fitting parameter as $0.044m_e$ for the bulk. Using Onsager's relation, the frequency F of the quantum oscillations can be directly related to the extremal cross-sectional areas S_k of the Fermi surface, as $F = 2\pi e S_k/h$. As the Fermi surface of Cd_3As_2 is to good approximation spherical, $S_k = \pi k_F^2$, the Fermi wavevector can be estimated as $k_F = (4\pi e F/h)^{1/2} \approx 3.33 \times 10^7 \text{ m}^{-1}$. The separation between the nodes of 0.8 nm^{-1} as reported in ref. 19 substantially exceeds the experimentally determined k_F . The Fermi velocity for our microstructure follows as $v_F = \hbar k_F/m^* \approx 8.76 \times 10^5 \text{ m s}^{-1}$. The Dirac equation now leads to an estimate of the Fermi energy above the Dirac point as $E_F = \hbar v_F k_F \approx 192 \text{ meV}$.

Additional complications for the analysis of the Shubnikov–de Haas effect arise from the quasi-ballistic transport in our clean microstructures that are thinner than the bulk mean free path. At very low fields, a peak in the magnetoresistance with a broad maximum followed by a strong negative magnetoresistance is observed before the bulk magnetoresistance is eventually recovered at high

fields (Fig. 3a). The quantum oscillations emerge on top of this background. We note that this field configuration is transverse and thus no negative magnetoresistance associated with the chiral anomaly is expected. Instead this enhancement of scattering in ultra-pure systems is a semi-classical effect arising from the diffuse scattering of otherwise ballistic electrons at the boundaries of strongly confined microstructures³⁸. It is well-studied in ultra-pure hydrodynamic systems, such as the viscous flow of ³He through capillaries (the Knudsen effect) or in high quality, geometrically confined conductors (Gurzhi flow) such as semiconductor heterostructures and clean metal whiskers³⁹. The resistance maximum at the Knudsen peak occurs at maximal boundary scattering of the bent electron trajectories, that is, at $2r_c = L$ up to a small numerical factor⁴⁰, where $r_c = \hbar k_F / (eB)$ is the cyclotron radius. As a result, the position of the maximum is expected to shift to higher fields as the sample thickness is decreased, as is observed in the Cd₃As₂ microstructures (Fig. 3a). The Knudsen effect is a direct consequence of ballistic electron motion between the opposite surfaces and thus is direct evidence that the thickness of the studied microstructures is indeed comparable to the bulk mean free path.

Non-adiabatic corrections. The positions of the Landau levels associated with F_S systematically deviate from the usual $1/B$ periodicity. Figure 4 shows the deviation of each maximum of the resistance oscillation, $1/B_n$, from its expected position $1/B_0$ for a usual $1/B$ periodicity (blue dashed line). At low fields, the oscillations indeed are found to be periodic in $1/B$, yet slightly but consistently deviate at higher fields. The deviations from periodicity are shown in Fig. 4b.

We find quantitative agreement between the observed deviation and the expectations for non-adiabatic corrections of the Weyl orbit (Fig. 4b). Yet deviations of similar magnitude from periodicity could also occur from spin-splitting of the surface state alone, without the presence of Weyl orbits. However, a comparison of the presented data with the predictions for spin-splitting shows that the observed peak positions are qualitatively different from expectations in the spin-splitting scenario, further highlighting the unusual character of the observed oscillations (Extended Data Fig. 4). In the presence of strong magnetic fields, the Fermi surface volume for spin-up and spin-down electrons changes due to the Zeeman energy $\Delta E = \frac{1}{2} \mu_B g \sigma_z H_z$, where the z axis is pointing along the magnetic field. This leads to an effective splitting of an initially spin-degenerate Fermi-surface at zero field (in the absence of spin-orbit coupling). In favourable cases, spin splitting can be directly observed as the appearance of two, well-separated peaks positioned symmetrically around the expected position in the absence of spin-splitting⁴¹. We do not observe split peaks in the surface state related oscillations. This could either be an indication of weak spin-splitting, or of a thermal or impurity broadened situation where direct splitting of the peaks is not observable despite strong spin-splitting. In the latter case, spin-splitting is known to modify the amplitude of the oscillations only, without affecting the phase of the oscillations³⁶. Therefore spin-splitting of a trivial surface state alone is unlikely to explain the present data.

Another possibility may be the formation of a surface state akin to those in a topological insulator (TI). In this case, the spin-momentum locking on the Dirac cone changes the spin-splitting behaviour and oscillation phase changes without any peak splitting being possible²⁸. The reason for this is the opening of a gap in the Dirac spectrum due to the time reversal symmetry breaking of the applied magnetic field. This gap due to the Zeeman energy modifies the Landau level spectrum for electrons (at positive band filling):

$$E_N = v_F(2ne\hbar H + (g\mu_B H / (2v_F))^2)^{1/2} \quad (2)$$

In the absence of spin-splitting ($g = 0$), this reduces to the conventional square-root-dependence of Dirac systems. At low fields, that is, large Landau level index n , this correction is negligible, yet becomes important at higher fields closer to the quantum limit ($n = 0$). The resulting fields B_n , where a Landau level equals the Fermi energy E_F can be easily calculated using the material parameters self-consistently obtained from the quantum oscillation analysis (blue dots in Extended Data Fig. 4a) and are contrasted to the measured peak positions (red dots in Extended Data Fig. 4a).

Spin-splitting of a TI-like surface state is at odds with our observations for two reasons: (1) the expected deviation goes in the wrong direction and (2) the observed magnitude is incompatible with measured g -factors in Cd₃As₂. Spin-splitting would suggest that the fields B_n occur at lower fields compared to the purely periodic case, they are on the contrary observed at higher fields. The shift to lower fields in the case of spin-splitting can be easily understood from the surplus Zeeman energy adding to the field dependence of the Landau levels in the absence of spin splitting (see below). This shifts the levels up in energy, and thus they intersect the constant chemical potential at lower fields, hence the upwards bending of the fields B_n as shown in Extended Data Fig. 4b. In addition, a g -factor of 300 is required to obtain a shift of similar magnitude as observed by our experiment, which is an order of magnitude larger than the experimentally measured values⁴².

On the other hand, such a downwards deviation as well as its magnitude is expected for Weyl orbits arising from non-adiabatic corrections: at small fields, the quasiparticle evolving on one of the surface states may only enter the bulk state at the Dirac point, where the gap between the surface and bulk states vanishes. In strong magnetic fields, however, the quasiparticle may tunnel through this bandgap into the bulk state before reaching the Dirac point over a distance given by the magnetic length, $l_B = (\hbar / (eB))^{1/2}$. This shortens the effective path length on the surface (orange arrows in Fig. 1b) and thus leads to a non-adiabatic correction to the low-field surface frequency $F_{S,0}$ in strong fields given by

$$F_S(B) = F_{S,0} - 4\alpha F_{S,0} l_B / k_0 \quad (3)$$

where α represents a material dependent parameter encoding the tunnelling barrier between the surface and the bulk bands, and $F_{S,0}$ the field-independent surface frequency in the absence of strong non-adiabatic effects observed at low fields. The four tunnelling processes involved in each Weyl orbit thus lower the effective quantum oscillation frequency F_S in strong magnetic fields.

Including this non-adiabatic correction term effectively explains the observed deviation (Fig. 4b, red points) from the pure $1/B$ periodicity in high magnetic fields. A fit to the observed data (blue points) yields a coupling parameter $\alpha \approx 1.25$, in good agreement with the theoretical predicted value of order unity¹². This non-adiabatic correction to the $F_S(B)$ is also directly evident in the Fourier spectra: while the signature of the bulk frequency is a sharp, symmetrical peak centred at F_B , the surface frequency F_S is asymmetric with sizeable spectral weight shifted to lower frequencies. This again is a result of the effective lower frequency in high magnetic fields, and can be well reproduced by a simple calculation inserting the field-dependent frequency given by equation (3) into the Lifshitz-Kosevich formalism and Fourier-transform the resulting waveform (Fig. 4a). Unlike the spin-splitting driven phase modification, the non-adiabatic correction mechanism leads to a slowing-down of the oscillations at higher fields due to the effective reduction of flux enclosed in the quantum path that is cut short by the tunnelling process. This naturally leads to the down-bending observed in our experiment.

Thickness dependence of the oscillation phase. Equation (1) describes the expected width dependence of the oscillations arising from the Weyl orbit. Crucially, the thickness L does not affect the oscillation frequency, as the oscillation period Δ is independent of it: $\Delta = B_n^{-1} - B_{n-1}^{-1} = ek_0^{-1} \pi v_F / E_F$. In agreement with this expectation, no thickness-dependence of the frequency is observed. Instead it is expected to only affect the phase offset of the peaks versus $1/B$ and indeed the data show no thickness dependence to the oscillation frequency. However, experimentally detecting this phase shift is a challenge for four main reasons: (1) the surface- and bulk-frequencies are close to each other leading to strong beating which complicates the identification of the peak positions corresponding to the surface frequency, (2) the peak positions show an additional field dependence due to the non-adiabatic corrections, (3) the oscillatory signal is on top of a large nonlinear and thickness-dependent background magneto-resistance due to finite size effects which must be subtracted to find the peak positions, and (4) the relative amplitude of the surface oscillations compared to the bulk oscillations by itself is strongly thickness dependent, which causes a thickness dependence of the beating pattern. Combined, these uncertainties in fitting the peak positions are larger than the separation between peaks, and therefore from the present data it is impossible to make a quantitative claim about the systematic dependence of the peaks. These difficulties are naturally avoided in the triangular geometry and thickness dependence experiments. By design, the triangular sample averages over parts of different effective width, leading to a cancellation of the surface signal as expected from equation (1).

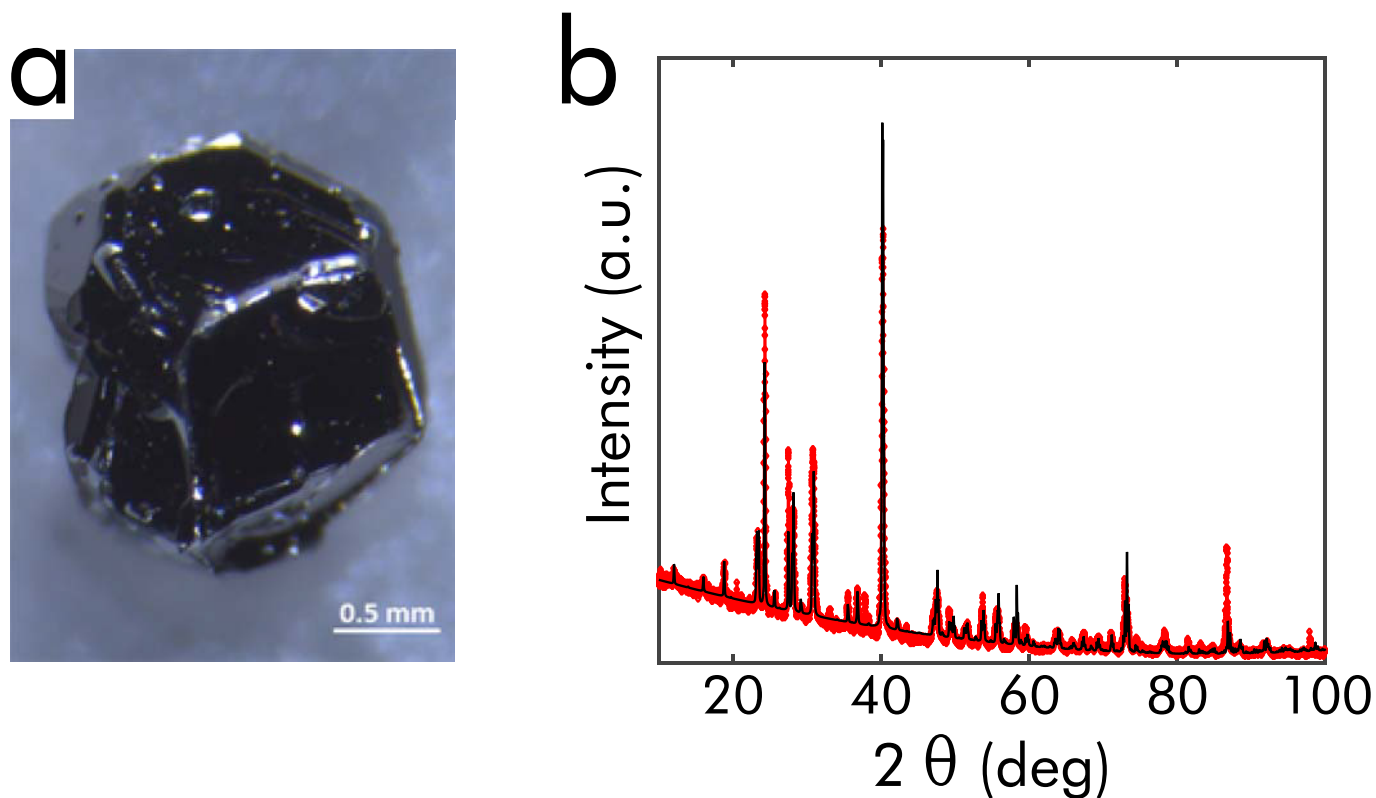
Upper field limit for surface state oscillations. Reference 12 predicted that quantum oscillations associated with the Weyl orbit persist only up to a maximum field that depends on sample thickness, $B_{\text{sat}} \approx k_0 / (eL)$. By repeating this analysis, we show that the thickness dependence is actually more involved. Instead, the surface arc oscillations persist up to a field that depends in a complicated non-monotonic fashion on film thickness, but which is at least as large as the quantum limit associated with the k -space area enclosed by the Fermi arcs. To see this, note that from equation (1) above we can see that the smallest integer, $n = N$, for which equation (1) has a solution is given by: $N = \left\lceil \frac{k_F L}{\pi} - \gamma \right\rceil$, where $k_F = E_F / v_F$, and $\lceil x \rceil$ indicates the closest integer to x that is larger than x . N represents the index of the last quantum level associated with the Weyl orbits that can be pushed across the Fermi energy. The field at which level N crosses the Fermi energy, that is, the upper limit field beyond which Weyl surface-arc oscillations cease, is then: $B_{\text{sat}} = \frac{k_F k_0}{\pi} \frac{1}{N - \left(\frac{k_F L}{\pi} - \gamma \right)}$. The first factor in B_{sat} is precisely the $1/B$ frequency of the surface state, $F_S = 56$ T. The second factor depends strongly on L in a complicated oscillatory fashion, but is typically of order 1 (though possibly much

larger when $(k_F L/\pi) - \gamma$ is accidentally close to an integer value). Hence, we see that, in contrast to the claims of ref. 12, the upper limit field for observing Weyl surface arc oscillations is comparable to the quantum limit of the surface state, and is largely independent of sample thickness. In the present measurements, B_{sat} is expected to be at least ~ 60 T, and is hence unobservable in any non-pulsed field experiment.

Raw oscillation data from the 150 nm device. As quantum oscillations are phenomena most easily understood in Fourier space, we limit the discussion in the main text on the analysis of FFT spectra. Fourier transforms, however, can be misleading and thus it is very important that the main results are apparent in the raw data themselves. This is evidently the case: Extended Data Figs 5 and 6 show the oscillatory magnetoresistance after subtracting a second-order polynomial to remove the non-oscillatory components. The same notation of angle is used as in the main text. Scans were performed in 5° steps around the surface state orientations (0° , 180°) and in 10° steps elsewhere for the 150 nm device, and in 5° steps for the triangle. No averaging or symmetrization of the data between field sweeps of opposite polarity was performed, that is, between sweeps with 180° angle difference. From the data set Extended Data Fig. 5, Fig. 2b showing the surface character of the oscillations was constructed. The main result of the manuscript can be seen as follows: for in-plane fields, that is, 90° , 270° , a single frequency corresponding to the bulk frequency is observed. As discussed in the main text, both the observation of only one single frequency as well as its amplitude is in agreement with all other published data of quantum oscillations on Cd_3As_2 . For fields perpendicular to the surface, that is, at 0° , 180° , the spectrum is profoundly different. An additional oscillation of higher frequency is mixed with the bulk frequency, leading to a beating pattern. The second frequency gradually disappears as the field is rotated away from 0° , and is identified with the surface state oscillation from the FFT spectrum as described in the main text. Fig. 5 showing the differences between the rectangular and triangular devices was constructed from the raw data set shown in Extended Data Fig. 6. The main points can also be well seen here without resorting to Fourier transforms: despite its identical surface and cross-sectional area, the triangle does not show any sign of a second frequency at any angle, as can be seen from the angle-independent

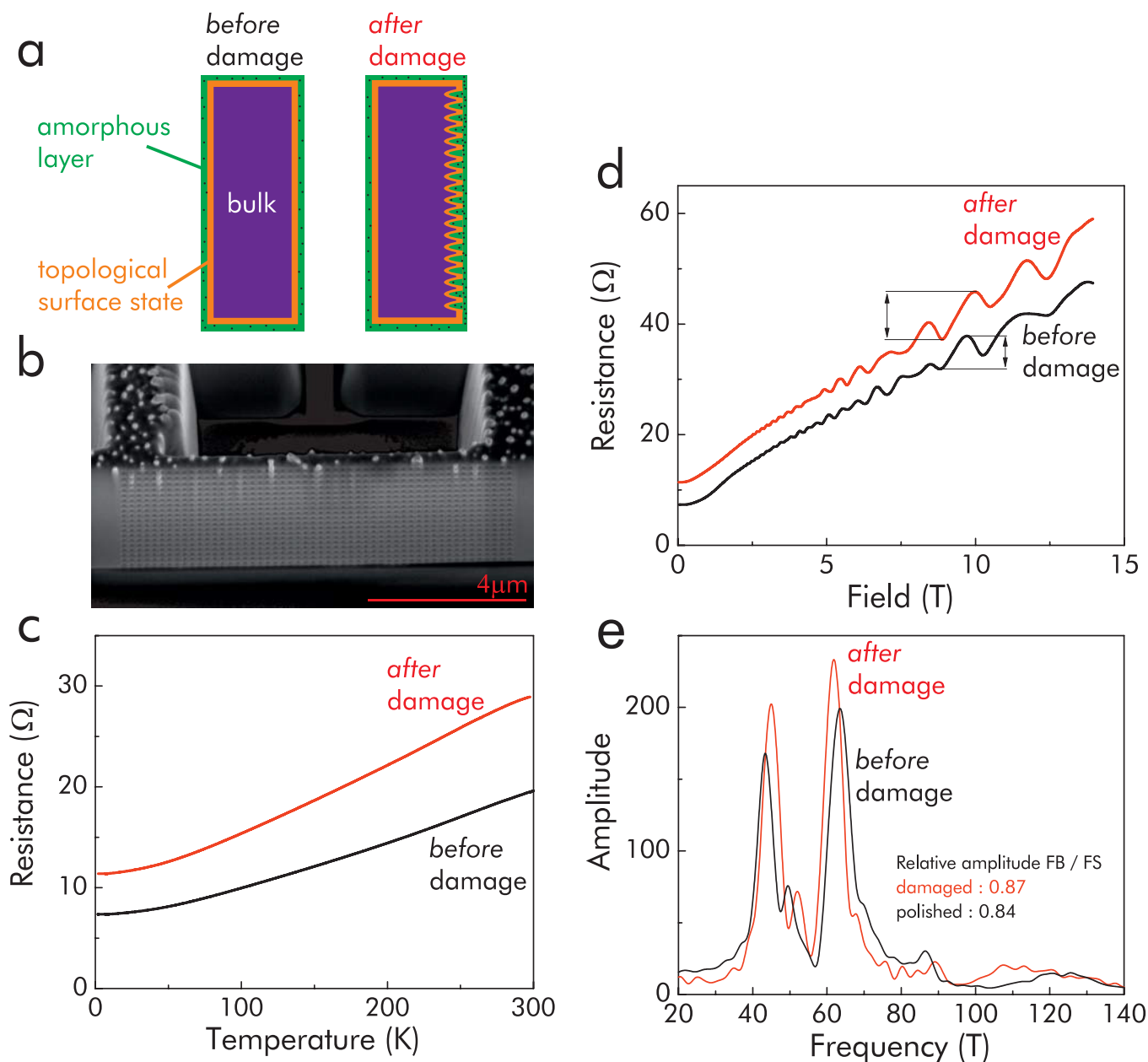
positions of the quantum oscillation peaks. In the rectangular case, however, strong beating appears when the field is perpendicular to the surface, signalling the second frequency. Also, the characteristic angle dependence of surface states as presented in Fig. 2 can be clearly seen when one traces the maxima from the surface to higher angles.

29. Moll, P. J. W. *et al.* Transition from slow Abrikosov to fast moving Josephson vortices in iron pnictide superconductors. *Nat. Mater.* **12**, 134–138 (2012).
30. Moll, P. J. W., Zhu, X., Cheng, P., Wen, H.-H. & Batlogg, B. Intrinsic Josephson junctions in the iron-based multi-band superconductor $(\text{V}_2\text{Sr}_4\text{O}_6)\text{Fe}_2\text{As}_2$. *Nat. Phys.* **10**, 644–647 (2014).
31. Moll, P. J. W. *et al.* Field induced density wave in the heavy fermion compound CeRhIn_5 . *Nat. Commun.* **6**, 6663 (2015).
32. Jaroszynski, J. *et al.* Upper critical fields and thermally-activated transport of $\text{Nd}(\text{O}_{0.7}\text{F}_{0.3})\text{FeAs}$ single crystal. *Phys. Rev. B* **78**, 174523 (2008).
33. Ooi, S., Mochiku, T. & Hirata, K. Periodic oscillations of Josephson-vortex flow resistance in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+y}$. *Phys. Rev. Lett.* **89**, 247002 (2002).
34. Moll, P. J. W. *et al.* High magnetic-field scales and critical currents in $\text{SmFeAs}(\text{O}, \text{F})$ crystals. *Nat. Mater.* **9**, 628–633 (2010).
35. Ziegler, J. F. SRIM-2003. *Nucl. Instrum. Methods Phys. Res. B* **220**, 1027–1036 (2004).
36. Shoenberg, D. *Magnetic Oscillations in Metals* (Cambridge Univ. Press, 1984).
37. Wray, L. A. *et al.* How robust the topological properties of Bi_2Se_3 surface are: a topological insulator surface under strong Coulomb, magnetic and disorder perturbations. *Nat. Phys.* **7**, 32–37 (2011).
38. Beenakker, C. W. J. & van Houten, H. Quantum transport in semiconductor nanostructures. *Solid State Phys.* **44**, 1–228 (1991).
39. de Jong, M. & Molenkamp, L. Hydrodynamic electron flow in high-mobility wires. *Phys. Rev. B* **51**, 13389–13402 (1995).
40. Thornton, T. J., Roukes, M. L., Scherer, A. & Van De Gaag, B. P. Boundary scattering in quantum wires. *Phys. Rev. Lett.* **63**, 2128–2131 (1989).
41. Tokumoto, M. *et al.* Direct observation of spin-splitting of the Shubnikov-de Haas oscillations in a quasi-two-dimensional organic conductor (BEDT-TTF) $_2\text{KHg}(\text{SCN})_4$. *J. Phys. Soc. Jpn* **59**, 2324–2327 (1990).
42. Wallace, P. R. Electronic g-factor in Cd_3As_2 . *Phys. Status Solidi* **92**, 49–55 (1979).
43. Ali, M. N. *et al.* The crystal and electronic structures of Cd_3As_2 , the three-dimensional electronic analogue of graphene. *Inorg. Chem.* **53**, 4062–4067 (2014).



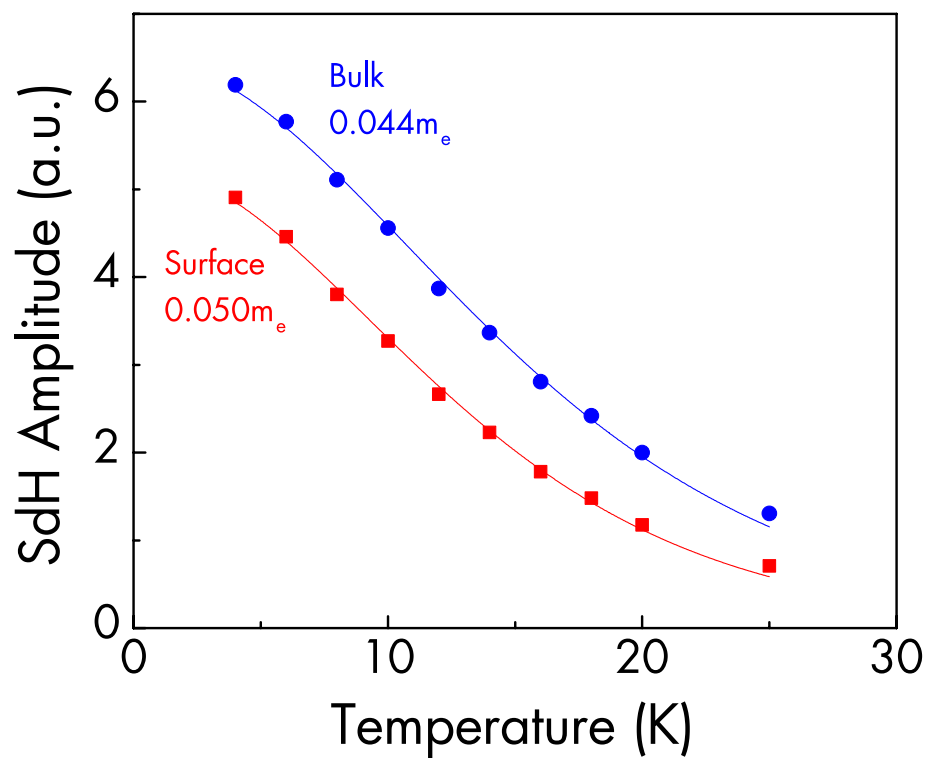
Extended Data Figure 1 | Synthesis and characterization of Cd_3As_2 crystals. **a**, A typical faceted Cd_3As_2 single crystal, and **b**, the powder X-ray diffraction (PXRD) spectrum. Cd_3As_2 single crystals were produced using the flux growth technique with a 5:1 ratio of Cd flux to Cd_3As_2 . Elemental Cd and As were placed into an aluminium oxide crucible with a quartz wool plug and sealed into a quartz ampoule under vacuum. The ampoule was heated to 825 °C and held at that temperature for two days

to ensure a fully homogenized mixture. It was then cooled at a rate of 6 °C per hour to 425 °C, where it was centrifuged to remove excess Cd. Bulk crystals were found trapped in the quartz wool following this procedure, and were confirmed to be Cd_3As_2 by both PXRD and single-crystal X-ray diffraction. The PXRD data (red trace) can be well fitted by the $I4_1/acd$ low-temperature phase of Cd_3As_2 (ref. 43) (black trace), and no parasitic phases were observed.



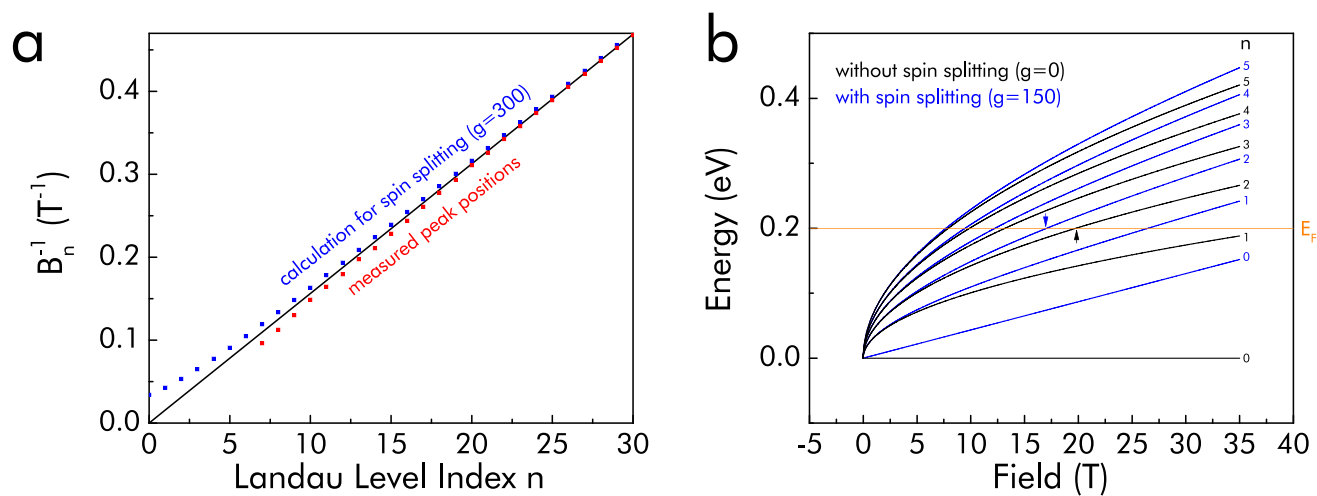
Extended Data Figure 2 | Surface quantum oscillations are resistant to intentional surface damage. **a**, Sketch of the device cross-section before and after the irradiation. Three possible current paths exist in these devices: the crystal bulk (purple), the topological surface states (orange), and the amorphous FIB-induced damage shell (green). **b**, Device after the heavy irradiation damage. The dimples due to the beam centre impact

are well visible across the whole device. The polished back side remained undisturbed by this procedure. **c**, Comparison of the same device with pristine surfaces and after damaging the surface. The resistance increases after irradiation, as expected for the increased scattering. **d**, Magnetoresistance of the Cd_3As_2 microstructure before and after irradiation; and **e**, the Fourier components of the quantum oscillations.



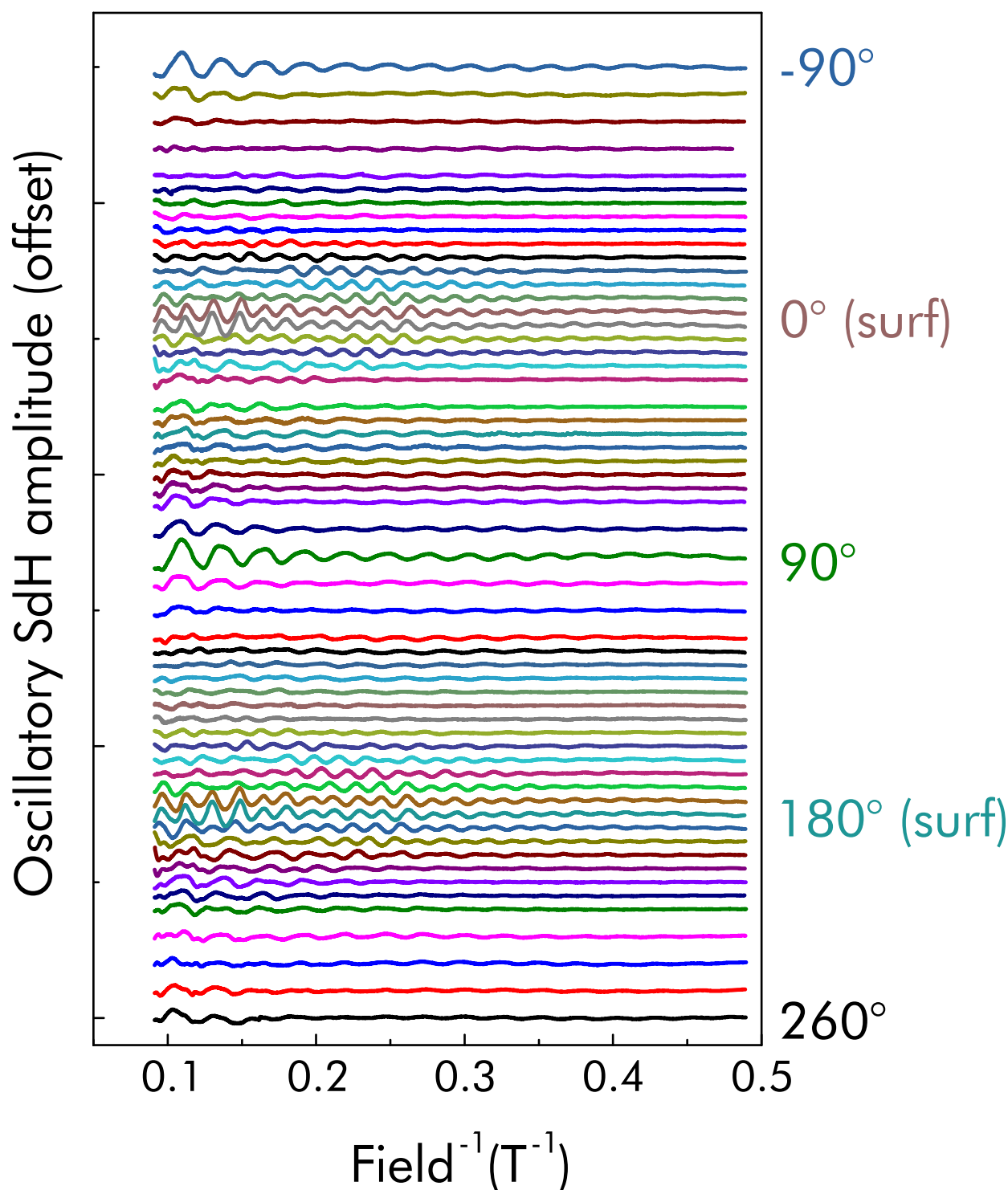
Extended Data Figure 3 | Effective mass analysis of bulk and surface oscillations. The temperature dependence of the Shubnikov–de Haas (SdH) amplitudes in the bulk (blue) and at the surface (red) follow the usual Lifshitz–Kosevich behaviour. The fits to the data (lines) yield an

effective electron mass in the bulk of $0.044m_e$, in good agreement with previously reported values²⁶. The surface state appears slightly heavier ($0.05m_e$).



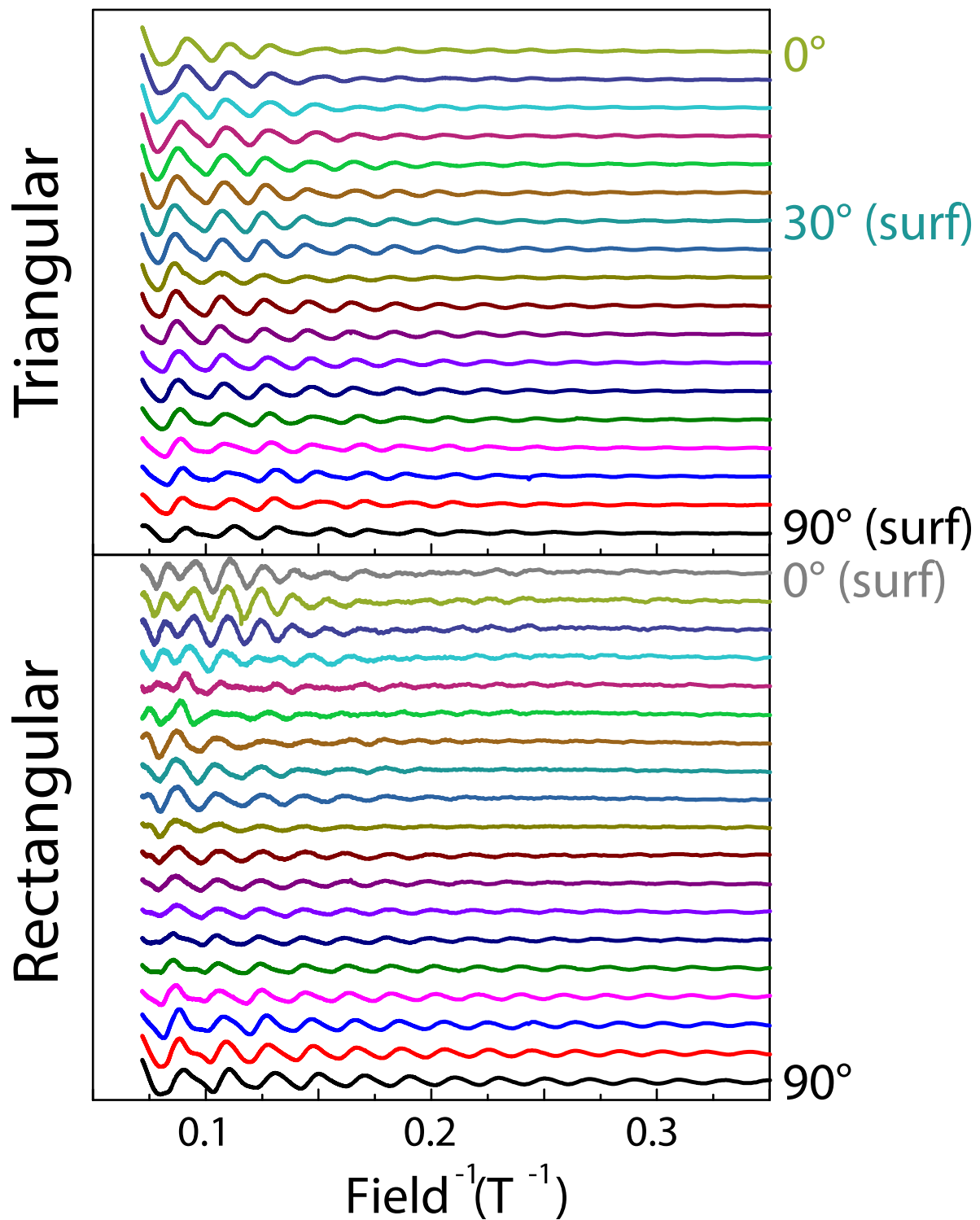
Extended Data Figure 4 | Spin splitting. **a**, Positions of the n th Landau levels, B_n , calculated for spin-splitting (blue) and experimentally observed values (red). The observed deviation is opposite to the expectation of spin-splitting, yet qualitatively and quantitatively consistent with the

non-adiabatic corrections of the Weyl orbit process. **b**, Calculated Landau level energy spectrum for Dirac systems with and without spin-splitting as described by equation (2).



Extended Data Figure 5 | Raw oscillatory signal of the 150 nm device. Shown are Shubnikov–de Haas oscillations of the 150-nm-wide rectangular device, the smallest microstructure fabricated in this study, as a function of $1/\text{field}$, for different angles between the sample and the magnetic field. Following the notation of the main manuscript, 0° and 180° correspond to a field perpendicular to the surface of the device ('surf'). The separation between curves along y represents the field angle

on a linear scale. Close to the perpendicular field configurations, traces at 5° angle increment were taken, and at 10° increment elsewhere. The appearance of the surface frequency can be well seen in the raw data as strong beating appears around 0° . Also, the characteristic $\cos(\theta)^{-1}$ angle dependence of a surface frequency can be easily seen by following the peak positions to higher angles.



Extended Data Figure 6 | Raw oscillatory signal of the triangular and rectangular devices. Shown are Shubnikov-de Haas oscillations of the triangular (top) and rectangular (bottom) devices, as a function of $1/\text{field}$ for different angles between the sample and the magnetic field. Traces were

taken at 5° angle increment. Those angles where the field is perpendicular to a surface are marked by 'surf' for both the rectangular and the triangular devices. While the rectangular device shows the characteristic beating, the peak positions in the triangle remain at the same fields.

Self-assembly of graphene ribbons by spontaneous self-tearing and peeling from a substrate

James Annett^{1,2,3} & Graham L. W. Cross^{1,2,3}

Graphene and related two-dimensional materials have shown unusual and exceptional mechanical properties^{1–3}, with similarities to origami-like paper folding^{4,5} and kirigami-like cutting^{6,7} demonstrated. For paper analogues, a critical difference between macroscopic sheets and a two-dimensional solid is the molecular scale of the thin dimension of the latter, allowing the thermal activation of considerable out-of-plane motion. So far thermal activity has been shown to produce local wrinkles in a free graphene sheet that help in theoretically understanding its stability⁸, for example, and give rise to unexpected long-range bending stiffness⁶. Here we show that thermal activation can have a more marked effect on the behaviour of two-dimensional solids, leading to spontaneous and self-driven sliding, tearing and peeling from a substrate on scales approaching the macroscopic. We demonstrate that scalable nanoimprint-style contact techniques can nucleate and direct the parallel self-assembly of graphene ribbons of controlled shape in ambient conditions. We interpret our observations through a simple fracture-mechanics model that shows how thermodynamic forces drive the formation of the graphene–graphene interface in lieu of substrate contact with sufficient strength to peel and tear multilayer graphene sheets. Our results show how weak physical surface forces can be harnessed and focused by simple folded configurations of graphene to tear the strongest covalent bond. This effect may hold promise for the patterning and mechanical actuating of devices based on two-dimensional materials.

In this Letter we report how the thermal activation of the thin dimension of a two-dimensional solid allows it to undergo self-tearing, sliding and folding, which result in substantial mass transport. In Fig. 1 we show spontaneous, large-scale motion of graphene. A single-layer sheet was mechanically exfoliated onto a 300 nm thermal oxide layer on a silicon wafer (optical micrograph Fig. 1a) and then heated to 150 °C on an open hot plate for 13 d. The subsequent optical micrograph of the flake in Fig. 1b shows a large, folded-over flap formed by a 5 µm horizontal tear in the top of the flake combined with a straight, vertical fold that is over 20 µm in length. A series of steps that might have led to the formation of this flap is proposed in the schematics of Fig. 1c–e: the process was initiated by the formation of a small, stable folded-over tab (Fig. 1c). We assume that this tab is too small to see in the optical micrograph of Fig. 1a, and was probably created as an unintended by-product of the original mechanical nucleation process. This embryonic structure then grew by sliding and peeling, and eventually also tearing, as it evolved from Fig. 1c–e.

To justify and explain this proposed formation sequence, in the lower half of Fig. 1 we demonstrate a deterministic mechanical method that allows the controllable nucleation of an embryonic fold such as that in the schematic in Fig. 1c, which then results in a spontaneous growth process. In Fig. 1f–l, a time series of atomic force microscopy (AFM) images reveals the formation of three elongated ribbons in bilayer graphene that decorate a triangular indentation crater of about 1 µm in size. The crater was formed by performing ambient, room-temperature

nanoindentation using a pyramidal diamond Berkovich tip with a load sufficient to pierce the sheet and permanently indent the substrate below. Critical to the success of the method was the presence of a small amplitude (~1 nm) lateral shearing oscillation imposed at high frequency throughout the indentation process. This fretting action was found to form small tabs of folded-over, adhered graphene—an example of which is shown in the inset of Fig. 1l. The fretting was necessary to realize growing ribbon structures (>70% yield with fretting versus 0% without, with over 100 ribbons produced in one-, two-, three- and four-layer graphene, see Fig. 1n for an example of high-yield production.) For the three ribbons nucleated in Fig. 1l, each had a starting width of about 0.6 µm and tapered inwards with a ~15° angle. The upper ribbon (1) grew to length of ~1.5 µm, whereas the lower left ribbon (2) encountered and grew over a subsheet defect before also stopping at ~1.5 µm in length. The ribbon on the right (3) halted at ~0.7 µm in length when it encountered the flake edge. A time-lapse movie of this 14 day process can be found in Supplementary Information.

We have produced ribbons with starting widths of ~300 nm to >2,000 nm by varying the indentation depth; some of these ribbons were observed grow to >5 µm in length. Heating the samples above room temperature (either globally on a hot plate or locally using a focused laser spot) increased the nucleation probability and generally accelerated and expanded growth, however, the majority of our ribbons were nucleated and grew at room temperature. Once underway, most of the length of a ribbon formed on a timescale shorter than AFM imaging (which is on the order of minutes), but then continued slowly, decelerating smoothly over any period of continued image capture that we used. Ribbon growth could be stochastic, sometimes impeded by obvious external defects or contaminants (see Extended Data Fig. 1) that were pushed aside or otherwise overcome to allow rapid growth to resume. Although many ribbons grew in a symmetric fashion, deviations were observed. Collisions with defects, contamination, flake edges and other ribbons—or varying conditions such as temperature during growth—produced asymmetric, transient fold bulging with variable, asymmetric or even negative tapering. This led to asymmetry in ribbon shape and rotation of the growth direction.

Analysis of the strips peeled from adhered macroscopic films has shown how tearing forces and strip tapering angles depend on adhesion, fracture toughness and elastic strain energy within the fold^{9,10}. In our case, peeling of the graphene strips is clearly not the result of an external force¹¹, but instead, we propose, the result of a thermodynamic mechanism of exchanging graphene–substrate area with graphene–graphene area. We refer to this as the interfacial force, arising from the difference in the energy release rates of peeling and healing along the entire width of the ribbon at the attaching fold. Our system thus consists of a closing adhesive crack that drives the tearing of two covalent cracks, sliding of the ribbon head and peeling from the substrate.

For sufficiently wide ribbons, we find that the net exchange of interfacial energy is sufficient to produce motion forwards against the forces of sliding and tearing resistance. The stability of this folded configuration

¹Centre for Research on Adaptive Nanostructures and Nanodevices, Trinity College, Dublin 2, Ireland. ²Advanced Materials and BioEngineering Research, Trinity College, Dublin 2, Ireland. ³School of Physics, Trinity College, Dublin 2, Ireland.

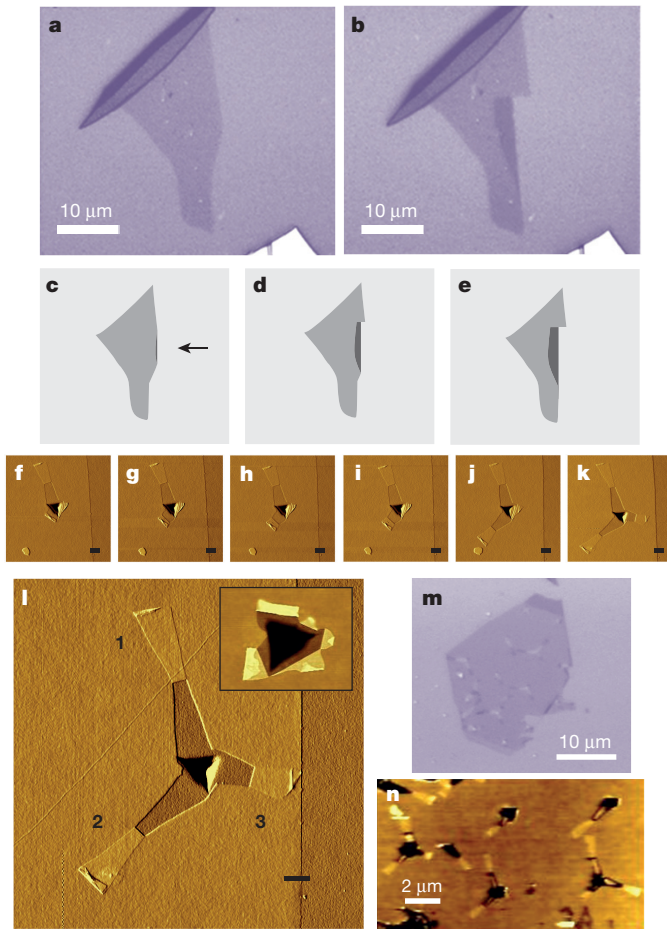


Figure 1 | Self-assembly of graphene flaps and ribbons. **a–e**, A self-folded graphene flap formed by spontaneous folding, sliding, peeling and tearing of a single-layer sheet adhered to a silicon oxide substrate. **a, b**, Optical micrographs taken before (**a**) and after (**b**) the spontaneous formation of a 20- μm -long folded flap by tearing back from the right-hand edge of the flake. The system was heated to 150 °C in ambient conditions for 13 d between **a** and **b**. The final area of the measured flap is approximately 50 μm^2 . **c–e**, A proposed series of steps leading to the formation of the tab is shown, with an initial nucleating fold indicated by the black arrow in **c**, and progressive growth by sliding and tearing in **d**, to the final position in **e**. **f–n**, Directed graphene ribbon self-assembly by nanoindentation. **f–l**, AFM image sequence of room-temperature, ambient condition growth of three ribbons nucleated from single fretted nanoindentation contact (similar to the inset in **l**) in bilayer graphene over a period of 14 days. Panels **f, g, h, i** and **j** show progressive stages in the growth of ribbon 2. Ribbon 2 grows over a sub-sheet defect in **k** and **l**, whereas ribbon 3 starts growth between **j** and **k** and ceases growth prematurely at the flake edge in **l**, avoiding growth onto the silicon oxide substrate. A time lapse movie of this growth is provided in Supplementary Information. Scale bars, 500 nm. **m**, Optical micrograph of high-yield, multiple ribbon growth events nucleated by fretted contact in a bilayer graphene flake. **n**, AFM image of the central portion of **m** at higher magnification.

is enabled by a length of graphene–graphene surface interaction that overcomes its own folding strain energy U_{fold} , similar to the structure of collapsed nanotubes^{12,13} or a self-adhered graphene ‘tennis racket’^{4,14}, but altered by adherence to the rigid flat substrate^{15–17}. We expect the strain energy contained within the fold to influence the fracture path, reducing the width of the fold⁹.

In the following we present an analysis of the forces acting during the final stages of ribbon growth to give physical insight into the overall formation process. An AFM image of a ribbon formed from a trilayer graphene flake and with no (or imperceptibly slow) growth is shown in Fig. 2c, along with height profiles (Fig. 2a, b) extracted from the image,

indicating the presence of a folded ridge with a protruding height of about 0.4 nm and a width of less than ~ 20 nm (which is the lateral resolution of our image). From a nucleated width w_0 of $\sim 0.8 \mu\text{m}$, the ribbon is formed by tearing mirror-symmetric, straight fracture paths, each of which tapers inwards with taper half-angle θ of $\sim 6^\circ$ with respect to the growth axis, reaching a final length l of $\sim 2.1 \mu\text{m}$ and a final folded ridge width w of $\sim 0.4 \mu\text{m}$. The ribbon head position h is also indicated. In the absence of contaminants, this form of ribbon growth is typical, exhibiting symmetric, inward tapering at a small, constant tapering angle with a highly symmetric, tightly folded and straight ridge terminated by two sharply defined tearing points. Invoking a Griffith-style elastic fracture analysis¹⁸, we assume that the host graphene sheet that contains the ribbon shown with area A_{ribbon} has total area A_{sheet} and write the internal energy of the system as

$$U = U_{\text{fold}} + 2\lambda c + \gamma_1(A_{\text{sheet}} - A_{\text{ribbon}}) + \gamma_2 A_{\text{ribbon}} \quad (1)$$

where λ is the rupture energy per unit of length required to tear the sheet (which ignores the orientation effects of crystallography), γ_1 is the adhesion energy of the sheet to the substrate, γ_2 is the adhesion energy of the sheet to itself and c is the path length of the two tears (assumed to be identical) that define the ribbon shape.

In general, for a two-dimensional material with a given number of layers and bending stiffness D , the strain energy U_{fold} will depend on the un-adhered excess ribbon length $\zeta = 2l - h$ contained in the fold as well as w (ref. 9), expressed explicitly as $U_{\text{fold}}(\zeta, w) = wD \int_0^\zeta \kappa^2(s) ds$ where $\kappa(s)$ is the curvature of the sheet at path position s in the folded region. However, similar to self-adhered folding in free graphene strips¹⁴, we expect in our system that beyond a minimum ribbon length, the cross-sectional fold shape (Fig. 2c, lower inset) becomes constant (that is, ζ becomes constant) and $U_{\text{fold}} = U_{\text{fold}}(w)$ only. Assuming for the moment that no friction is associated with the sliding of the adhered ribbon over its host sheet, we compute the fracture energy release rate¹⁸ of our system for a simultaneous increment δc of both tears under our interfacial force to be

$$\frac{\delta U}{\delta c} = \frac{\delta U_{\text{fold}}}{\delta w} \frac{\delta w}{\delta c} - (\gamma_2 - \gamma_1)w \frac{\delta l}{\delta c} + 2\lambda \quad (2)$$

Under our quasi-static, end-of-growth conditions, we have a balance of forces acting at each crack tip and the energy release rate must be zero¹⁸. Noting that $(\delta l / \delta c) = \cos\theta$, $(\delta w / \delta c) = -2\sin\theta$ and letting $S = \int_0^\zeta \kappa^2(s) ds$ remain constant, we divide the advancing and resisting effects equally between the two tears and from equation (2) can write the forces acting at each crack tip as

$$F_{\text{advance}} = \frac{\gamma_2 w \cos\theta}{2} + DS \sin\theta = \frac{\gamma_1 w \cos\theta}{2} + \lambda = F_{\text{resist}} \quad (3)$$

as shown in the upper schematic of Fig. 2c.

Projecting the forces derived in equation (3) along the direction of ribbon propagation, we find at each crack tip

$$\frac{(\gamma_2 - \gamma_1)w}{2} = \lambda \cos\theta \quad (4)$$

Experimentally, the dominant direction for fracture growth is found to be at right angles to the fold (that is, we have small θ), here seen to be driven by the net interface tension (difference between healing and peeling) force $(\gamma_2 - \gamma_1)w$ overcoming most of the tearing resistance. Along the fold direction (that is, orthogonal to the ribbon propagation direction) we find that $DS = \lambda \sin\theta$ where an energetically favourable tendency to reduce the fold energy⁹ manifests as a small, inward taper force working against a small tearing resistance component. These projections are also represented in the upper schematic of Fig. 2c. The maximum energy release rate criterion of fracture mechanics dictates that the tear will take the path of minimum force¹⁸, found by minimizing

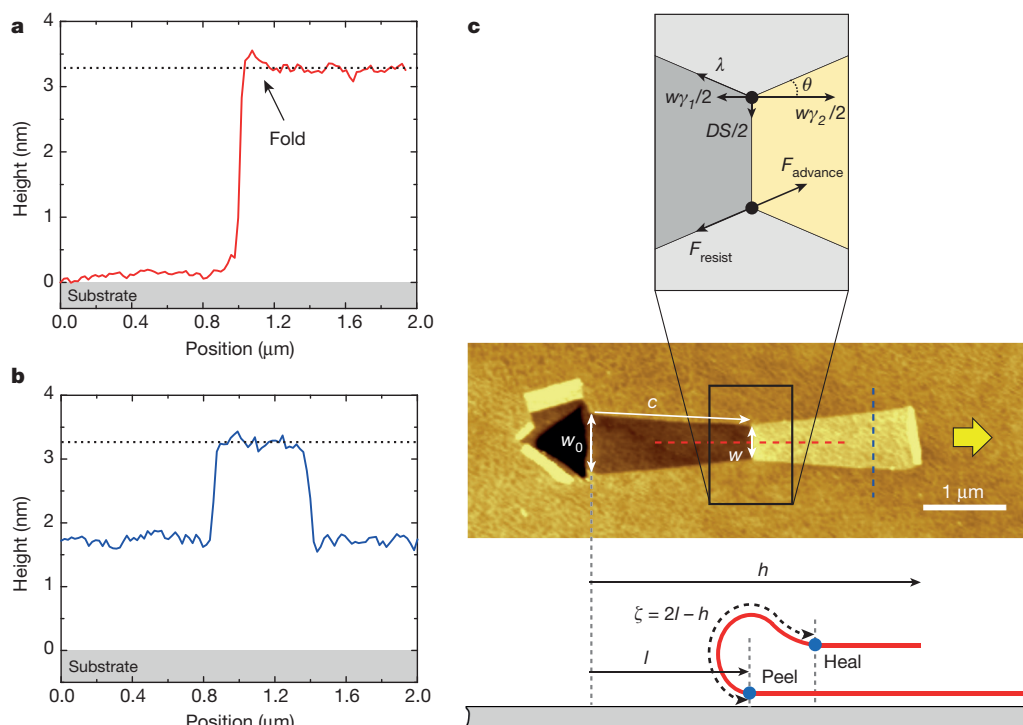


Figure 2 | Fracture mechanics analysis of quasi-static ribbon growth. **a, b**, Cross-sectional profiles extracted from **c** (along the dashed red and blue lines) of a typical self-assembled ribbon undergoing slow, quasi-static growth with symmetric, constant tapering in trilayer graphene. The dashed black lines indicate the top of the ribbon. **c**, AFM image of the ribbon with the geometric parameters of ribbon length l , head position h , nucleation fold width w_0 and final fold width w labelled. The upper schematic shows the plan view of two equivalent representations of the forces acting at each tearing crack tip (black dots) as well as the definition

of the taper half-angle θ . Ribbon growth occurs due to an interfacial healing crack with force magnitude $w\gamma_2/2$ focused at each tear tip, which overcomes the peeling and tearing resistances of $w\gamma_1/2$ and λ respectively. Inward taper occurs due to a tendency to reduce strain energy in the fold, manifesting as force $DS/2$ at each tip. The lower schematic is a side view of the ribbon showing a (highly zoomed) view of the fold cross-section with excess length $\zeta = 2l - h$ (dashed black curve between surface peel and heal points, indicated by blue dots). Ribbon growth direction is to the right, as indicated by the yellow arrow.

equation (3) with respect to θ . Performing this operation, we find the taper half-angle to be related to the ratio of the driving forces as

$$\tan\theta = \frac{2DS}{(\gamma_2 - \gamma_1)w} \quad (5)$$

The significance of small, constant taper angles observed for a majority of the ribbons is highlighted by equation (5). First, the presence of substantial frictional sliding should manifest as an extra force resisting growth in equation (4) and add to the denominator of equation (5). Classically, we expect friction to increase directly with ribbon area¹⁹, thus the constant taper angle we observe for ribbons without externally impeded growth (that is, not affected by contamination) implies that friction is of negligible magnitude compared to the interfacial force. It seems that fold-constrained (that is, fixed-direction) sliding of graphene on supported graphene maintains low friction in ambient conditions as ribbons grow well above the micrometre scale. Sliding graphitic basal interfaces with incommensurate orientation have been shown to exhibit superlubricity on nanometre²⁰ to micrometre²¹ scales, even in ambient conditions²². We have confirmed the incommensurate stacking of our ribbons using scanning Raman spectroscopy (see Extended Data Figs 2 and 3). This is consistent with observations in other systems of weak scaling of friction with area due to mismatched lattice structure at an atomically flat interface¹⁹.

In the absence of friction, the steady attenuation of the driving interfacial force due to tapering as indicated by equation (4) suggests why ribbon growth stops at a finite width. In this case, ribbons nucleated with different starting widths should taper equivalently and terminate with the same final width under similar conditions. Figure 3a shows three bilayer graphene ribbons with parallel growth direction, nucleated with different starting widths in close proximity by nanoindentation.

This configuration gives a high probability that the ribbons will share the same surface adhesion conditions and crystallography of a single grain in the graphene. In Fig. 3b we overlay the traces of the ribbons to show that they share a common taper angle $2\theta \approx 12^\circ$ and stopping width $w \approx 290$ nm despite the fact that each has a very different size and presumably different friction. To further investigate this, we have measured the shape of a population of 15 contamination-free ribbons that exhibit a constant taper angle in one-, two- and three-layer graphene. The final widths of our controlled population are plotted in Fig. 3c, which shows increasing final widths with an increasing numbers of layers. Assuming that the tearing resistance increases with the number of layers, this supports the assertion that ribbon growth is terminated by the decreasing interfacial force at a threshold width where it is insufficient to tear the graphene lattice. In principle, the stopping width and taper angle can be predicted from the sheet properties in equations (4) and (5), however the quantitative scaling behaviour of the graphene bending modulus^{11,16} and adhesion^{23–26} with layer number remains unclear, and compounded by our lack of knowledge of the local crystallography (which is expected to influence λ), we only report our measurements here.

Overall, our analysis suggests a picture of a well-adhered but easily sliding ribbon head, with the newly torn ribbon length drawn through a fold with an invariant cross-section that is held in constant strain per unit width by the balance of adhesion forces. Although we expect the angle of the inward tapering of the ribbon to slowly increase due to the taper-induced decrease of the interfacial force in equation (5), this is generally not apparent before ribbon growth ceases. Figure 3a also suggests that our ribbon self-assembly may be exploited to perform positive- or negative-tone lithography. Large-area scale-up of multiple nucleation sites should be readily achievable with multiple-tip nanoimprint-style stamps. The formation time of unimpeded

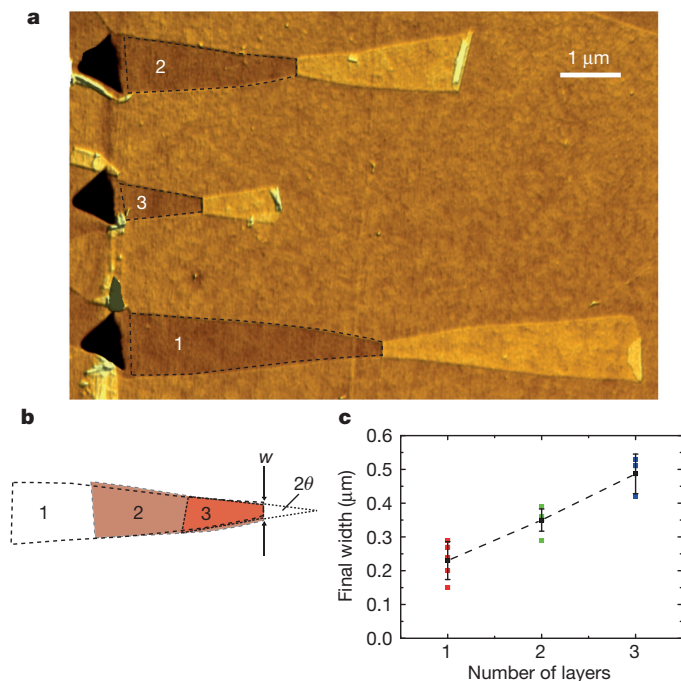


Figure 3 | Geometry of ribbon growth. **a**, Simultaneous growth of three ribbons (labelled 1–3) in bilayer graphene with different starting widths. **b**, Numbered trace overlays demonstrating that all three ribbons have a similar tapering angle $2\theta \approx 12^\circ$ and identical final stopping width $w \approx 290$ nm. **c**, Plot of w versus the number of layers for a population of 15 ribbons including single-, bi- and trilayer graphene (shown in red, green and blue, respectively). The black squares indicate the mean final width for each layer count and the error bars show one standard deviation (the black dashed line is shown to guide the eye to the trend). The increase in the final width of the ribbons with layer number supports a picture of ribbon growth being arrested by the tearing resistance.

self-assembly is faster than AFM scanning can determine, and may indeed take place in a fraction of a second even at room temperature—similar to microscale graphitic pillar pull-in²².

In the slow, terminal regime of ribbon growth we expect sheet tearing to be governed by lattice trapping, where a force insufficient to break the carbon bonds in the sheet is augmented by thermal activation¹⁸. In Fig. 4 an AFM image sequence of room-temperature bilayer ribbon growth is shown with an approximately constant taper angle, captured over a period of 3.5 h. The semi-logarithmic plot in Fig. 4 presents over two decades of smooth, decelerating ribbon growth velocity data fitted to the form

$$v = \nu_0 a_0 \left[\exp \left(-\frac{\lambda a_0}{k_B T} \right) \sinh \left(\frac{\alpha (\Delta\gamma - \beta) w}{k_B T} \right) \right] \quad (6)$$

where k_B is Boltzmann's constant, T is temperature, the natural lattice frequency $\nu_0 = k_B T / h \approx 6 \times 10^{12} \text{ s}^{-1}$, h is Planck's constant, a_0 is the spacing between the bonds along the crack path, α is the activation length, $\Delta\gamma = \gamma_2 - \gamma_1$ is the net interfacial energy difference and β embodies any active resistance terms beyond the substrate peeling and tearing (such as the friction of the ribbon head sliding or the internal fold inter-layer sliding), which we have assumed to also scale proportionally with ribbon width w . Equation (6) (derived in Methods with numerical analysis shown in Extended Data Fig. 4) relates the single-bond breaking, rate-activated velocity to ribbon width, where the latter quantity is assumed to be proportional to the crack-driving interfacial energy release rate. From the fit we extract a bond-rupture energy density of $\lambda \approx 30 \text{ eV nm}^{-1}$, which is broadly compatible with simulations of the graphene tearing resistance^{11,27}. If we further assume negligible dissipative processes (that is, $\beta = 0$), we find that the minimum difference in interface energies required to drive forward ribbon assembly in our

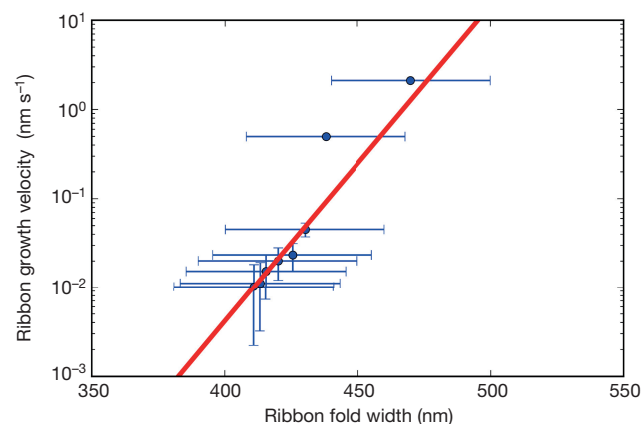


Figure 4 | Kinetics of self-assembling graphene ribbons. Semi-logarithmic plot of the velocity of ribbon growth in bilayer graphene versus the width of the ribbon fold w near the end of growth (the slow, quasi-static limit), showing evidence for stress-activated sheet rupture and frictionless sliding driven by interfacial adhesion forces that are proportional to the ribbon width. Data was extracted from the ribbon growing in the lower right of the sequence of AFM images shown below the plot. Unlabelled scale bars, 2 μm. The error bars in the plot indicate the standard deviation of the width and length measurements from the images.

sample is $\sim 0.02 \text{ eV nm}^{-2}$, which we note is only a small fraction of the adhesion energies that have been determined for graphene–solid interfaces, which are on the order of 1 eV nm^{-2} (refs 23–26). Multiplying this surface energy difference by the terminal width of ~ 350 nm we compute an interfacial driving force of approximately 1 nN. This value is slightly lower than the expected carbon bond strength of a few nanonewtons, consistent with fracture propagation at a sub-critical force for bond breaking. From this we infer that a slight increase in the adhesive affinity from one surface to another can drive self-assembly, provided that there is free sliding.

Over 20 years ago, it was envisioned that graphene could be folded and cut into useful forms as a kind of nanoscale origami⁵. Our results suggest a much richer potential for two-dimensional materials: they can behave as a self-animated sheet that folds, tears and slides over itself in ambient conditions. Spontaneous ribbon growth offers a way to investigate the combined effects of adhesion, folding, sliding and tearing of two-dimensional sheets in a self-consistent way. The weakness of sliding friction indicates that the growth of sufficiently wide ribbons could in principle reach very large scales suitable for wafer-scale transfer processes, for example. There is the potential to control and scale the behaviour as a lithography to pattern and assemble two-dimensional materials, including incommensurate hetero-stacks of dissimilar two-dimensional materials that may organize in novel ways depending on the strength of their interactions²⁸. Surface functionalization may allow the manipulation of self-assembly by heat, light or other controllable external fields, enabling nanoelectromechanical systems of mechanically coupled sensors and actuators made up of two-dimensional materials. Fundamentally, the behaviour we have observed reveals a tendency for isolated two-dimensional matter to coalesce towards a lower-energy three-dimensional form if kinetic barriers are overcome. The constrained sliding and folding pathways required to approach equilibrium may allow new arrangements and usages of two-dimensional materials to be realized through self-assembly.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 October 2015; accepted 18 March 2016.

1. Rasool, H. I., Ophus, C., Klug, W. S., Zettl, A. & Gimzewski, J. K. Measurement of the intrinsic strength of crystalline and polycrystalline graphene. *Nat. Commun.* **4**, 2811 (2013).
2. Zhang, T., Li, X. & Gao, H. Fracture of graphene: a review. *Int. J. Fract.* **196**, 1–31 (2015).
3. Ovid'ko, I. A. Mechanical properties of graphene. *Rev. Adv. Mater. Sci.* **34**, 1–11 (2013).
4. Cranford, S., Sen, D. & Buehler, M. J. Meso-origami: folding multilayer graphene sheets. *Appl. Phys. Lett.* **95**, 123121 (2009).
5. Ebbesen, T. W. & Hiura, H. Graphene in 3-dimensions: towards graphite origami. *Adv. Mater.* **7**, 582–586 (1995).
6. Blees, M. K. *et al.* Graphene kirigami. *Nature* **524**, 204–207 (2015).
7. Castle, T. *et al.* Making the cut: lattice kirigami rules. *Phys. Rev. Lett.* **113**, 245502 (2014).
8. Fasolino, A., Los, J. H. & Katsnelson, M. I. Intrinsic ripples in graphene. *Nat. Mater.* **6**, 858–861 (2007).
9. Hamm, E., Reis, P., LeBlanc, M., Roman, B. & Cerda, E. Tearing as a test for mechanical characterization of thin adhesive films. *Nat. Mater.* **7**, 386–390 (2008).
10. Kendall, K. Thin-film peeling—the elastic term. *J. Phys. D* **8**, 1449 (1975).
11. Sen, D., Novoselov, K. S., Reis, P. M. & Buehler, M. J. Tearing graphene sheets from adhesive substrates produces tapered nanoribbons. *Small* **6**, 1108–1116 (2010).
12. Chopra, N. G. *et al.* Fully collapsed carbon nanotubes. *Nature* **337**, 135–138 (1995).
13. Mikata, Y. New and improved analytical solutions for the self-folding problem of carbon nanotubes. *J. Nanomech. Micromech.* **3**, 04013004 (2013).
14. Meng, X., Li, M., Kang, Z., Zhang, X. & Xiao, J. Mechanics of self-folding of single-layer graphene. *J. Phys. D* **46**, 055308 (2013).
15. Cox, B. J., Baowan, D., Bacsá, W. & Hill, J. M. Relating elasticity and graphene folding conformation. *RSC Adv.* **5**, 57515–57520 (2015).
16. Chen, X., Yi, C. & Ke, C. Bending stiffness and interlayer shear modulus of few-layer graphene. *Appl. Phys. Lett.* **106**, 101907 (2015).
17. Chen, X., Zhang, L., Zhao, Y., Wang, X. & Ke, C. Graphene folding on flat substrates. *J. Appl. Phys.* **116**, 164301 (2014).
18. Lawn, B. R. *Fracture of Brittle Solids* (Cambridge Univ. Press, 1993).
19. Dietzel, D., Feldmann, M., Schwarz, U. D., Fuchs, H. & Schirmeisen, A. Scaling laws of structural lubricity. *Phys. Rev. Lett.* **111**, 235502 (2013).
20. Feng, X., Kwon, S., Park, J. Y. & Salmeron, M. Superlubric sliding of graphene nanoflakes on graphene. *ACS Nano* **7**, 1718–1724 (2013).
21. Liu, Z. *et al.* Observation of microscale superlubricity in graphite. *Phys. Rev. Lett.* **108**, 205503 (2012).
22. Yang, J. *et al.* Observation of high-speed microscale superlubricity in graphite. *Phys. Rev. Lett.* **110**, 255504 (2013).
23. Li, P., You, Z. & Cui, T. Adhesion energy of few layer graphene characterized by atomic force microscope. *Sens. Actuator. A* **217**, 56–61 (2014).
24. Vahdat, A. S. & Cetinkaya, C. Adhesion energy characterization of monolayer graphene by vibrational spectroscopy. *J. Appl. Phys.* **114**, 143502 (2013).
25. He, Y., Chen, W. F., Yu, W. B., Ouyang, G. & Yang, G. W. Anomalous interface adhesion of graphene membranes. *Sci. Rep.* **3**, 2660 (2013).
26. Bunch, J. S. & Dunn, M. L. Adhesion mechanics of graphene membranes. *Solid State Commun.* **152**, 1359–1364 (2012).
27. Huang, X., Yang, H., van Duin, A. C. T., Hsia, K. J. & Zhang, S. Chemomechanics control of tearing paths in graphene. *Phys. Rev. B* **85**, 195453 (2012).
28. Slotman, G. J. *et al.* Effect of structural relaxation on the electronic structure of graphene on hexagonal boron nitride. *Phys. Rev. Lett.* **115**, 186801 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank H. Nolan, N. McEvoy and G. Duesberg for discussions and assistance with the Raman spectroscopy. We thank A. Fasolino, A. Schirmeisen, J. Coleman and P. McCormack for discussions and J. Sader and J. Pethica for critical reviews of the manuscript. We acknowledge financial support from the Science Foundation of Ireland (SFI) under CRANN CSET 08/CE/I1432 and PI 08/IN/I1932. We also recognize assistance from Horizon 2020 COST Action MP1303.

Author Contributions J.A. observed the phenomena and performed all experiments. J.A. and G.L.W.C. designed experiments, analysed the data and constructed the static and kinetic models. G.L.W.C. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.L.W.C. (graham.cross@tcd.ie).

METHODS

Graphene flakes were mechanically exfoliated from Kish bulk graphite samples by the scotch-tape method and deposited on Si wafers with 300 ± 5 nm oxide. The thinnest flakes were selected by optical microscopy (Zeiss Axio Imager) and mounted in a nanoindentation system (Fast Forward Devices). Constant strain rate indentation with simultaneous small amplitude lateral oscillation using a diamond Berkovich tip was performed on the graphene flakes and the residual impressions and self-assembled ribbons were measured by AFM (Asylum MFP 3D) using tapping mode silicon probes (Budget Sensors). The layer number of the graphene flakes and the stacking order of the ribbons were assessed via scanning Raman spectroscopy (WiTec Alpha 300R) to micrometre spatial resolutions.

Care was taken to avoid the presence of defects during the experiments. We have observed numerous occurrences of collisions between spontaneously growing ribbons and local defects including contaminants, other ribbons and flake steps and edges. In Extended Data Fig. 1 we show the effect of ribbon collision with small, observable contamination particles (composition unknown) through a series of AFM images that tracked ribbon shape and position as a function of time over 12 h. The images in Extended Data Fig. 1 also demonstrate that ribbons can do useful work to move contamination particles, as indicated by the dotted arrow. Growth speed is greatly reduced during these interaction periods, but then resumes a high speed once the particle is cleared. The force exerted by impeding objects often results in a deformation of the fold, as shown by the yellow arrow in the figure. This deformation disappears upon removal of the impediment and resumption of normal growth.

The relative intensities, peak shape and peak position from Raman spectra have been used to determine the thickness of graphene²⁸ and the stacking order between layers^{30,31}. The ratio of the G to 2D band intensities and the full width at half maximum (FWHM) of the 2D band peak were used to determine the stacking order between the de-wetted graphene structures and the underlying flake^{30,31}.

An example of our Raman analysis for a mechanically exfoliated graphene flake with no self-assembled ribbons present is shown in Extended Data Fig. 2. An optical image of a mechanically exfoliated trilayer graphene sheet with a thicker four-layer strip is indicated in the optical micrograph in Extended Data Fig. 2a. A map of the G band intensity of this flake is shown in Extended Data Fig. 2b, which is a signature of the Raman characteristics of commensurately stacked graphene with an increase in the G band intensity within the thicker four-layer area. In contrast to this, the map of the 2D band, Extended Data Fig. 2c, shows a slight decrease in the 2D band intensity within the thicker four-layer area. This slight decrease is due to the stepwise broadening of the 2D band with increasing thickness, which results in an effective decrease in the 2D band intensity. This in turn results in an increase in the ratios of the G to 2D bands with increasing thickness, which were measured to be 1.6 and 2 within the three-layer and four-layer areas respectively. The broadening of the 2D band with increasing thickness can be seen by the map of the FWHM shown in Extended Data Fig. 2d and also by the plot of the averaged spectra within the three- and four-layer areas, shown in Extended Data Fig. 2e, f.

In Extended Data Fig. 3 we performed the same Raman analysis as described above on the three-ribbon growth sample shown in Fig. 1. Extended Data Fig. 3a shows an optical micrograph of the edge of the bilayer flake, where the three folded-over ribbon structures each form a four-layer stack. The map of the G band intensity, shown in Extended Data Fig. 3b is characteristic of the typical Raman signature of graphene with a larger G band intensity within the thicker folded-over sections of graphene. In contrast to the Raman signature of commensurately stacked graphene shown in Extended Data Fig. 2, the thicker folded-over graphene sheets result in a dramatic rise in the 2D band intensity, as shown in Extended Data Fig. 3d. In addition, the map of the FWHM of the 2D band shows no broadening within the thicker four-layer folded sections. The plots of the average spectra, shown in Extended Data Fig. 3e, f, highlight the similarity in the 2D band shape and G:2D ratios within the bilayer and four-layer areas, measured to be 1.44 and 1.47, respectively. The analysis indicates that the de-wetted graphene structures are incommensurate with the underlying flake and have not rotated into commensurability with the substrate.

For the ribbon growth kinetics, we present here a brief derivation of the rate activation analysis we used for the analysis of the data presented in Fig. 4, for single-bond breaking processes via the measurement of ribbon length and width versus time.

Fracture propagation can be considered as a stress-modified, thermally activated process of bond breaking with frequency¹⁸

$$K = \nu_0 \left[\exp\left(-\frac{\Delta F_{\pm}}{k_B T}\right) - \exp\left(-\frac{\Delta F}{k_B T}\right) \right] \quad (7)$$

where k_B is Boltzmann's constant, T is temperature and the natural lattice frequency is given

$$\nu_0 = \frac{k_B T}{h} \quad (8)$$

where F_{\pm} are the molecular free energies of formation of a stress-activated adsorption complex from the reactants and products respectively. The stress dependence can be expressed as a Taylor expansion of the difference $G - R$ of the crack energy release rate G from crack resistance energy density R

$$\Delta F_{\pm} = \Delta F \mp \alpha(G - R) + \dots \quad (9)$$

where ΔF is a quiescent adsorption-desorption activation energy at $G - R = 0$ and

$$\alpha = \left(\frac{dF}{dG} \right) \quad (10)$$

is the activation length. To find the crack velocity, we write $\nu = Ka_0$ where a_0 is the spacing between the bonds along the crack path. This gives

$$\nu = \nu_0 a_0 \left[\exp\left(-\frac{\Delta F}{k_B T}\right) \sinh\left(\frac{\alpha(G - R)}{k_B T}\right) \right] \quad (11)$$

The point $G = R$ is the quiescent point in ν - G space, giving a threshold for crack extension or retraction when the driving and resisting forces on the crack tip are balanced. From the static force analysis in the main text, we identify the individual bond rupture energy $\Delta F = \lambda a_0$. From equation (3), under quasi-static (slow) crack growth conditions for a 2D crack in a graphene sheet the energy release rate

$$G = \frac{\gamma_2 w \cos \theta}{2} + D \sin \theta \approx \frac{\gamma_2 w}{2} \quad (12)$$

and the resistance

$$R = \frac{\gamma_1 w \cos \theta}{2} + \beta w \approx \frac{\gamma_1 w}{2} + \beta w \quad (13)$$

where we assume possible resistance terms beyond the substrate peeling (such as the friction of ribbon head sliding and fold interlayer sliding) to scale with w over the range of observation with a proportionality constant β . Thus

$$\nu = \nu_0 a_0 \left[\exp\left(-\frac{\lambda a_0}{k_B T}\right) \sinh\left(\frac{\alpha(\Delta \gamma - \beta)w}{k_B T}\right) \right] \quad (14)$$

Quantitatively,

$$\nu_0 = \frac{k_B T}{h} \approx \frac{1/40 \text{ eV}}{4.14 \times 10^{-15} \text{ eV s}^{-1}} \approx 6 \times 10^{12} \text{ s}^{-1} \quad (15)$$

In the case of the bilayer graphene sample analysed in Fig. 4,

$$\nu_0 = \nu_0 a_0 \approx 6 \times 10^{12} \text{ s}^{-1} \times \frac{0.15 \text{ nm}}{2} \approx 4.2 \times 10^{11} \text{ nm s}^{-1} \quad (16)$$

where we have taken a mean bond length of 0.15 nm for single-layer graphene and doubled the number of bonds encountered by the crack in a bilayer sheet.

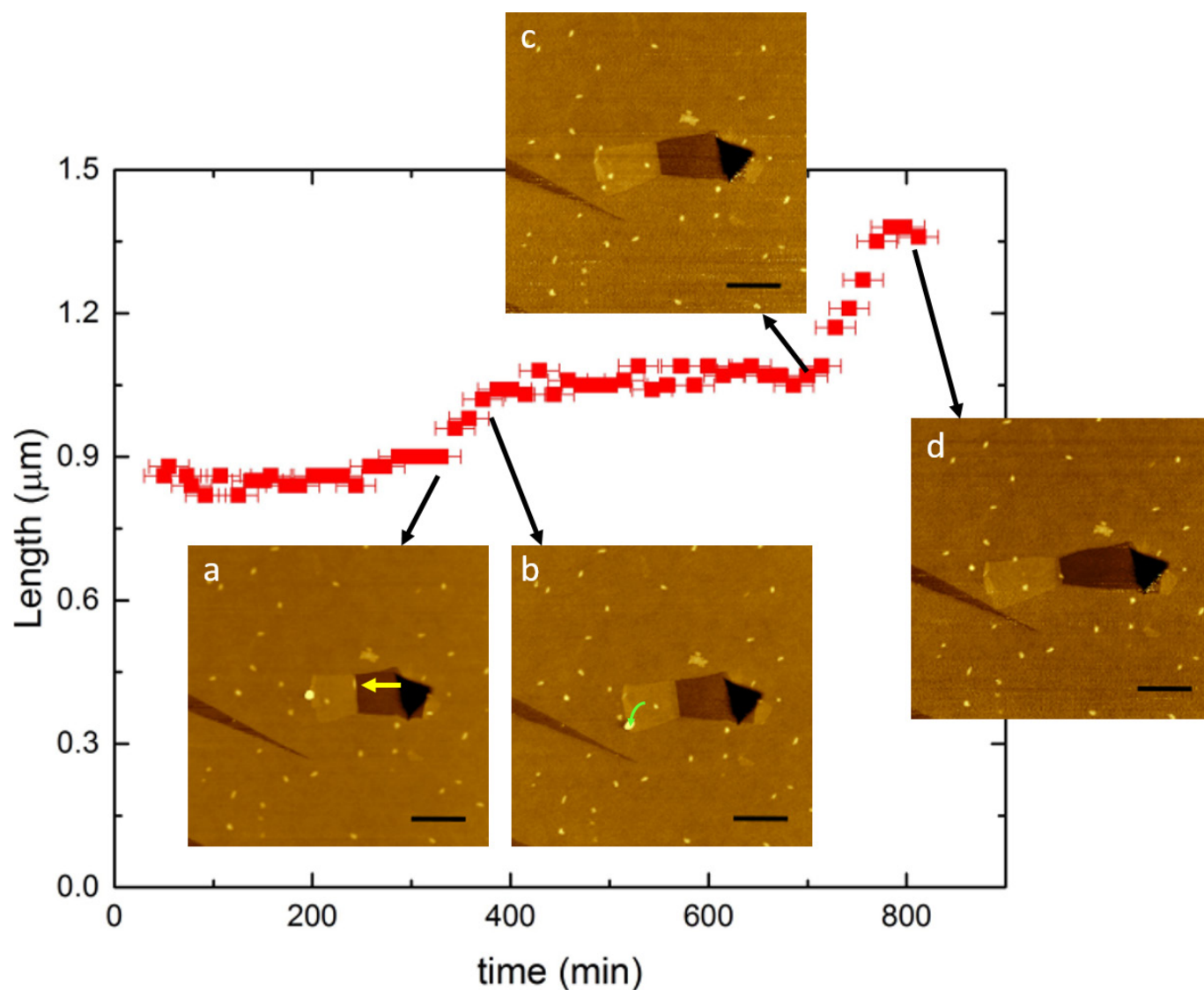
The direct measurement of the ribbon crack tip velocity versus time is hampered by a poor signal to noise ratio when taking a numerical derivative of our limited-precision AFM fold-position measurements. If we assume a constant taper angle over the range of ribbon growth measured, we can perform a numerical fitting procedure of the data to the path length versus the time of the integral of a logarithmic version of equation (14), which is a good approximation when the argument to the sinh function is greater than two, as we expect from the magnitude of the velocity we observe. For a fitted form to fold position l (giving approximately the path length for the small taper angle present) we use

$$l = a \log[b(t - t_0)] + l_0 \quad (17)$$

where (t_0, l_0) represent the point at which our quasi-static model begins to properly apply to the ribbon growth and a and b are fitting constants. We do not necessarily expect $(t_0, l_0) = (0 \text{ s}, 0 \text{ nm})$ due to possible unaccounted for high-speed growth physics, however we do assume that all of the measured growth data (on the order of minutes or more) has already reached the quasi-static regime. Extended Data Fig. 4 shows the results of fitting for the ribbon fold position and velocity to the measured data for the system presented in Fig. 4, as well as a similar procedure applied to the shrinking ribbon width w .

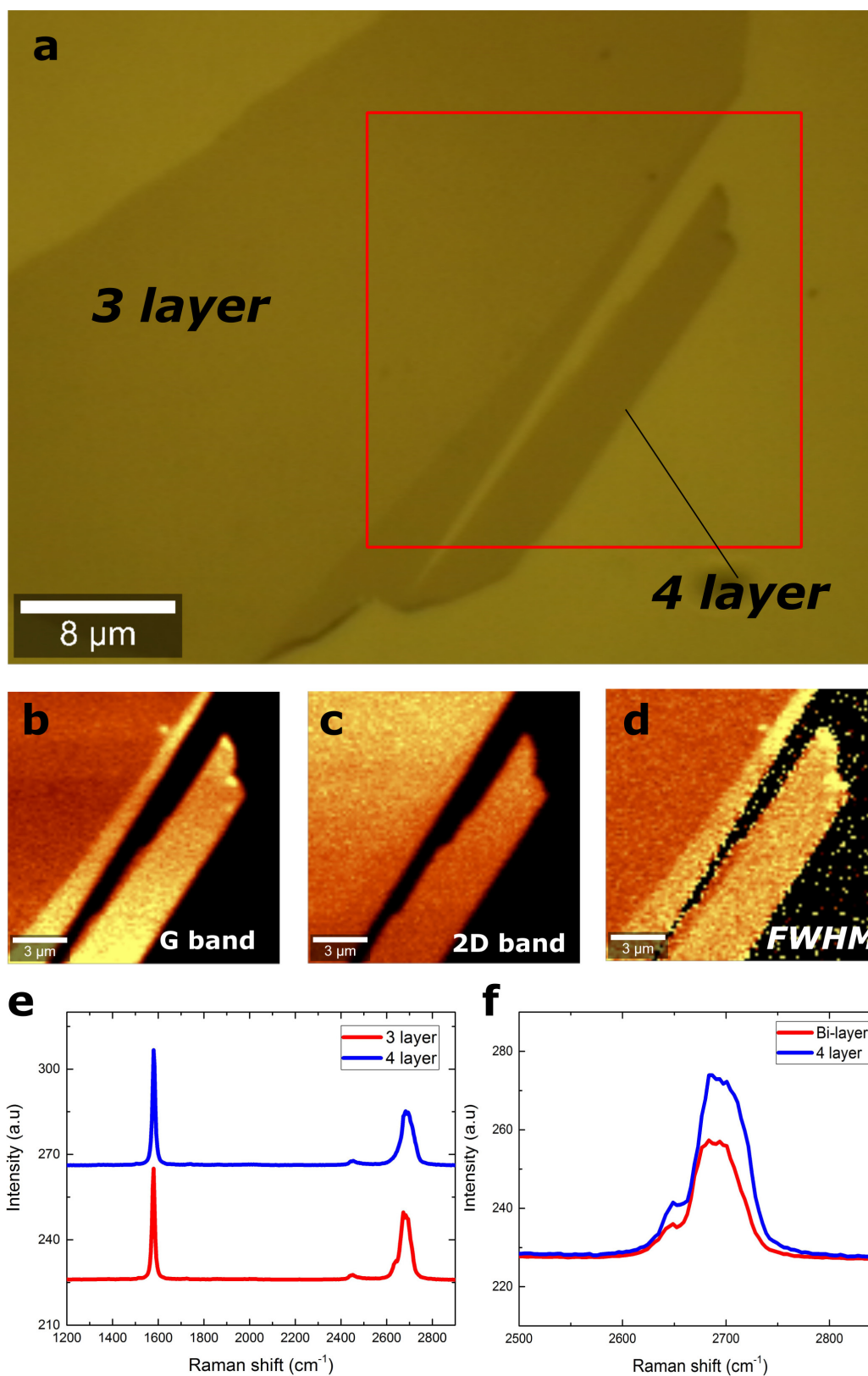
From the fit of velocity versus width in Fig. 4, we extract $\lambda a_0 = 3.0 \text{ eV}$ and $\alpha(\Delta\gamma - \beta) \approx 0.0021 \text{ eV nm}^{-1}$. Assuming a mean bond spacing of around 0.1 nm gives a bond-rupture energy density of $\lambda \approx 30 \text{ eV nm}^{-1}$, which is compatible with simulations of the graphene tearing resistance. The drag term β embodies the resistance due to finite sliding friction^{19,32} and other possible dissipative processes. Ignoring this term and assuming that the activation length α to be the bond spacing 0.1 nm , we find that $\Delta\gamma = 0.021 \text{ eV nm}^{-2}$.

29. Ferrari, A. C. *et al.* Raman spectrum of graphene and graphene layers. *Phys. Rev. Lett.* **97**, 187401 (2006).
30. Ni, Z., Wang, Y., Yu, T., You, Y. & Shen, Z. Reduction of Fermi velocity in folded graphene observed by resonance Raman spectroscopy. *Phys. Rev. B* **77**, 235403 (2008).
31. Hao, Y. *et al.* Probing layer number and stacking order of few-layer graphene by Raman spectroscopy. *Small* **6**, 195–200 (2010).
32. van den Ende, J. A., De Wijn, A. S. & Fasolino, A. The effect of temperature and velocity on superlubricity. *J. Phys. Condens. Matter* **24**, 445009 (2012).



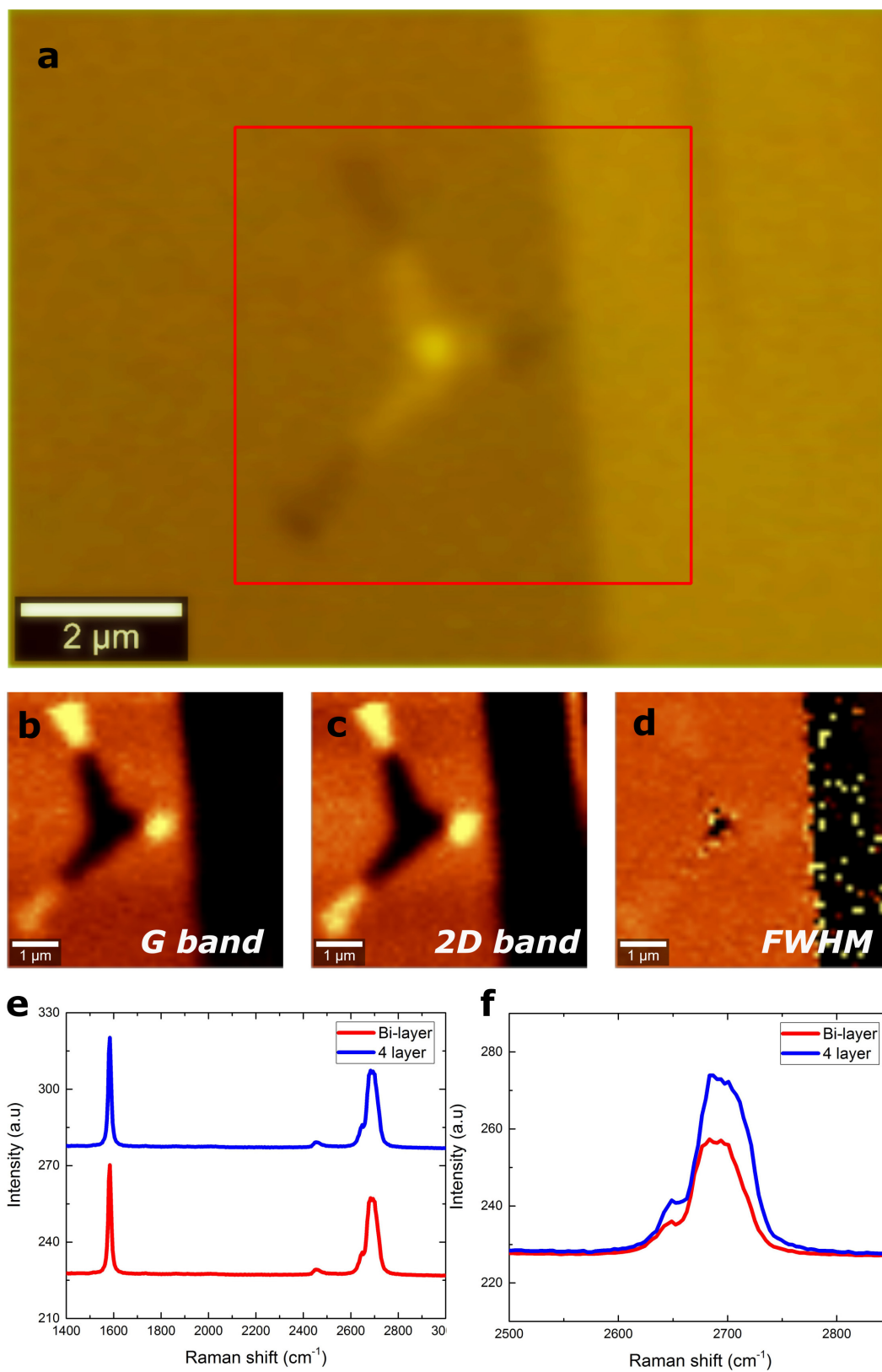
Extended Data Figure 1 | Ribbon growth and collision with visible contaminants. The growth pattern recorded by AFM imaging shows ribbon pinning and release as contamination defects are pushed aside (insets **a** and **b**), followed by rapid growth until next contaminant reached at point **c**. Final ribbon pinning occurs by encounter with sliver fracture in the flake that exposes the underlying substrate **d**. In **a**, a bulge in the

fold (yellow arrow) has been disappeared by **b** as the defect has moved to a lower position along the ribbon head indicated by the green arrow. Ribbon growth direction is to the left. Horizontal error bars indicate uncertainty in exact time of AFM image acquisition used to extract ribbon length values, whereas vertical error bars are omitted as the uncertainty range is too small to see at this scale.



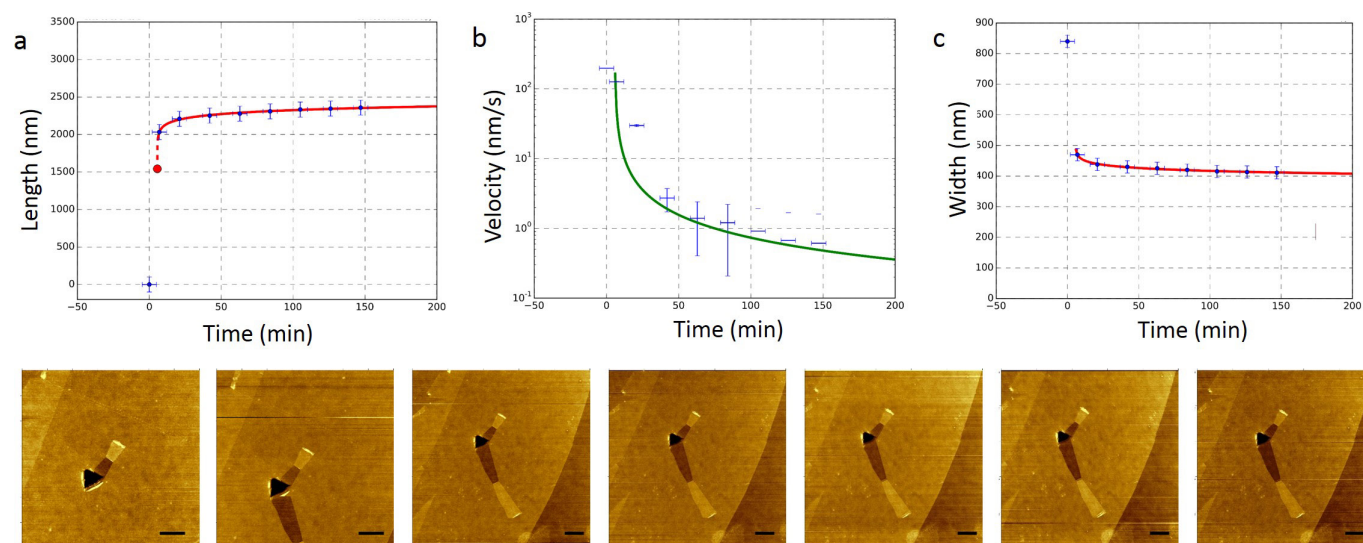
Extended Data Figure 2 | Scanning Raman analysis of commensurate stacked layers in a mechanically exfoliated graphene flake with no self-assembled structures present. **a**, Optical micrograph of a mechanically exfoliated graphene sheet. The red box indicates the area that was analysed, comprising three- and four-layer-thick graphene sheets.

b–d, Maps of the G band intensity (**b**), 2D band intensity (**c**) and the FWHM of the 2D band (**d**). **e**, Plot of the average Raman spectra within the three- and four-layer areas. **f**, Comparative plot of the 2D bands within the three- and four-layer areas.



Extended Data Figure 3 | Scanning Raman analysis of the self-assembled bilayer graphene ribbons shown in Fig. 11. a, Optical micrograph of the mechanically exfoliated bilayer graphene sheet. The red box indicates a single indentation from which the three folded ribbons

grew. **b–d,** Maps of the G band intensity (**b**), 2D band intensity (**c**) and the FWHM of the 2D band (**d**). **e,** Plot of the average Raman spectra within the folded ribbons (four-layer) and the bilayer sheet. **f,** Comparative plot of the 2D bands within the folded ribbons four-layer) and the bilayer sheet.



Extended Data Figure 4 | Fits to the ribbon head position. a–c, Ribbon length (a), velocity (b) and width (c) versus time for the self-assembling ribbon system presented in Fig. 4, as extracted from the sequence of AFM images shown below. See Methods for an explanation of the fitting functions. Scale bars, 1 μm .

Mid-ocean-ridge seismicity reveals extreme types of ocean lithosphere

Vera Schlindwein¹ & Florian Schmid¹

Along ultraslow-spreading ridges, where oceanic tectonic plates drift very slowly apart, conductive cooling is thought to limit mantle melting¹ and melt production has been inferred to be highly discontinuous^{2–4}. Along such spreading centres, long ridge sections without any igneous crust alternate with magmatic sections that host massive volcanoes capable of strong earthquakes⁵. Hence melt supply, lithospheric composition and tectonic structure seem to vary considerably along the axis of the slowest-spreading ridges⁶. However, owing to the lack of seismic data, the lithospheric structure of ultraslow ridges is poorly constrained. Here we describe the structure and accretion modes of two end-member types of oceanic lithosphere using a detailed seismicity survey along 390 kilometres of ultraslow-spreading ridge axis. We observe that amagmatic sections lack shallow seismicity in the upper 15 kilometres of the lithosphere, but unusually contain earthquakes down to depths of 35 kilometres. This observation implies a cold, thick lithosphere, with an upper aseismic zone that probably reflects substantial serpentinization. We find that regions of magmatic lithosphere thin dramatically under volcanic centres, and infer that the resulting topography of the lithosphere–asthenosphere boundary could allow along-axis melt flow, explaining the uneven crustal production at ultraslow-spreading ridges. The seismicity data indicate that alteration in ocean lithosphere may reach far deeper than previously thought, with important implications towards seafloor deformation and fluid circulation.

Mid-ocean ridges continuously produce new ocean lithosphere that consists of a layer of ocean crust (on average 6–8 km thick) underlain by mantle lithosphere⁷. The upper part of the mantle lithosphere is mechanically strong and brittle to temperatures of $650 \pm 100^\circ\text{C}$ (ref. 8) and is referred to here as elastic lithosphere. The base of the oceanic lithosphere is defined by an isotherm of $1,000\text{--}1,300^\circ\text{C}$ (ref. 9). Once the spreading rate drops below about 20 mm yr^{-1} conductive cooling reduces melt production and a different class of mid-ocean ridge forms that produces anomalous ocean lithosphere in 10%–20% of the world's oceans². Geological exploration of the poorly accessible ultraslow spreading ridges in the Arctic Ocean and the Southwest Indian Ocean revealed up to 100-km-long rift sections with mantle rocks exposed at the sea floor^{2–4}, indicating little mantle melt production. Between these sections pronounced volcanic centres receive more melt than the regional average and show an over-thickened crust³. Further observations that clearly distinguish ultraslow-spreading ridges from other mid-ocean ridges are the unexpectedly high incidence rate of hydrothermal plumes relative to the low magma budget¹⁰, a particular, smooth seafloor morphology⁶ and the potential to produce moment magnitude $M > 6$ earthquakes⁵. However, the lithospheric structure of ultraslow-spreading ridges is little known. Classic terms that describe the ocean lithosphere, such as 'crust' and 'Moho', lose their meaning in the absence of magmatism¹¹. Active-source seismic imaging of this anomalous lithosphere has been conducted only recently at the most accessible parts of ultraslow-spreading ridges and is unable to characterize lithospheric structure beyond its shallowest domains¹².

Records of local seismicity have greatly advanced our understanding of active spreading processes, the lithospheric structure and the thermal regime of most spreading ridges¹³ but no such data exist for ultraslow-spreading ridges. Following a 10-day feasibility study near Logachev Seamount on Knipovich Ridge¹⁴ (Fig. 1, site 2), we undertook the first concerted effort to compare the seismicity of contrasting magmatic and amagmatic sections of ultraslow spreading ridges. We recorded local earthquakes on one of the most extensive amagmatic sections of ultraslow spreading ridges, the oblique supersegment on the Southwest Indian Ridge (SWIR) (Fig. 1, site 1), with eight ocean bottom seismometers (OBSs) for a period of 11 months in 2012–2013 (ref. 15). The data set is unique because the long-term deployment of OBSs in the stormy Furious Fifties (the area between 50°S and 60°S , which is prone to strong winds), where only few research vessels can operate, has not been attempted before. A similar OBS network was installed simultaneously at a magmatic section, the Segment-8 volcanic centre on the eastern SWIR (Fig. 1, site 3) where an episode of teleseismic earthquake swarms occurred between 1996 and 2003 (ref. 5).

We determined the hypocentre locations of all of the recorded microearthquakes (see Methods). Figure 1 and Extended Data Figs 1–3 show a distinct image of hypocentre depths along 390 km of ultraslow-spreading ridges representing the prevalent background seismicity at local magnitudes of $M_l = 0.5\text{--}3$. The unprecedented clarity and consistency of the seismicity throughout the three representative sections allows us to state three key observations that expand previous knowledge of mid-ocean-ridge seismicity with important inferences for lithospheric structure.

(1) Earthquakes reach maximum depths of 35 km below the sea floor in areas with peridotite exposure, surpassing by far the deepest known mid-ocean-ridge earthquakes at about 15 km depth¹⁶.

(2) The clear lower boundary of seismicity, which provides constraints on the thickness of the axial elastic lithosphere (see Methods), varies dramatically along-axis, thinning by up to 15 km under volcanoes or sites of basalt exposure.

(3) Regions in the upper lithosphere are entirely aseismic. They occur mainly in peridotite-dominated ridge sections and extend to 15 km depth (Fig. 1; site 1, 40–75 km along the profile).

Melt flow along an undulating permeability boundary near the lithosphere–asthenosphere boundary towards magmatic centres is a key hypothesis in ultraslow ridge research that has been put forward by several authors^{17–19} to account for the pronounced variations in the crustal thickness, seafloor rock composition and tectonic structure observed along ultraslow spreading ridges. Our data show that the base of the elastic lithosphere as imaged by maximum hypocentre depths varies dramatically along-axis, being shallowest under volcanic centres (Fig. 1). If we assume that the permeability boundary near the lithosphere–asthenosphere boundary displays the same topography (see Methods) we can estimate the along-axis extent of the catchment areas where melts can flow upslope towards magmatic centres. These areas are at least 60–120 km in length and are hence consistent with the typical

¹Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Am Alten Hafen 26, 27568 Bremerhaven, Germany.

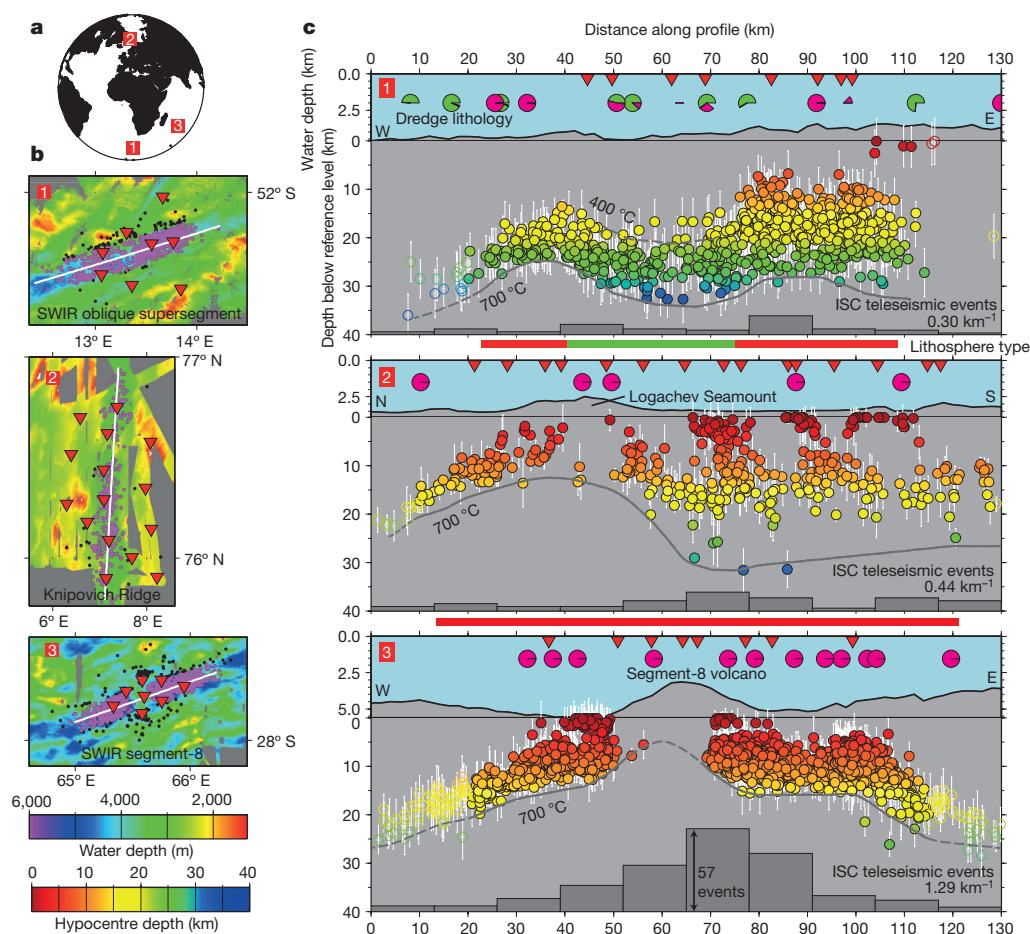


Figure 1 | Along-axis seismicity of ultraslow-spreading ridges. **a**, Study sites 1–3 on Knipovich Ridge and the SWIR. **b**, Maps of the study areas. White lines indicate the positions of transects in **c**. Red triangles are the OBS positions. Black dots mark earthquake locations, magenta dots are earthquakes that are projected onto the transects in **c**. All high-resolution bathymetry data are from RV *Polarstern* except SWIR segment-8 (ref. 6). **c**, Along-axis hypocentre depths (indicated by the colour coding, open circles are for those beyond the network) with error bars showing the 95% confidence level. The temperature regime is indicated by the estimated isotherms (grey lines, dashed where uncertain). The grey histograms indicate the numbers of teleseismic earthquakes from the Reviewed Bulletin of the International Seismological Centre (ISC; <http://www.isc.ac.uk/iscbulletin/search/bulletin/#reviewed>) along the ridge axis⁵ with the average event rates per kilometre of the rift axis labelled. The bathymetry is exaggerated vertically by a factor of 2.5 above the reference level (defined by the deepest OBS) marked by a thin horizontal line. The dredge lithology is shown in the pie charts (site 1 data from ref. 2, data for sites 2 and 3 are from <http://www.earthchem.org/petdb>): green, peridotite; magenta, basalt. The lithosphere type (green for amagmatic, red for magmatic) is indicated underneath the profiles where the OBS network provides sufficient constraint.

segment lengths of ultraslow-spreading ridges of about 100 km^{17,20}. Our data thus provide geophysical support for the hypothesis of melt focusing over segment-scale distances.

Peridotite may alter to mechanically weak phyllosilicates such as serpentinite (lizardite) in fractures that are penetrated by seawater up to temperatures of about 400 °C (refs 21 and 22). Volume fractions of only 10% serpentinite can drastically reduce the strength of the lithosphere and may lead to strain localization in distinct aseismic shear zones²³. We propose that the extensive aseismic regions, observed in peridotite-dominated, amagmatic ridge sections (Fig. 1; site 1, 40–75 km along the profile), result from serpentinization of the upper lithosphere either along distinct, deep-reaching shear zones that concentrate strain or through pervasive alteration of at least 10% of the mantle rocks. In this interpretation, the observed low seismic velocities at these depth levels (Extended Data Fig. 1b) reflect partial serpentinization²³. The onset of seismicity, and hence of brittle faulting below, occurs at a depth where temperatures of roughly 400 °C (for the estimate see Methods) are reached and serpentinite becomes unstable. The upper limit of seismicity in amagmatic regions may therefore image the serpentinization front at previously unknown depths of up to 15 km. This implies in turn that fluid circulation extends to these depths to alter the mantle rocks, exploiting major shear zones or a network of microfractures²⁴.

The adjacent magmatic regions, however, exhibit brittle deformation at depths of 8–15 km in the upper mantle or even throughout the lithosphere (Fig. 1; site 1, 75–115 km along the profile, site 2). Serpentinization in these mantle domains is apparently less

pronounced and cannot effectively reduce the shear strength, so these regions of the mantle behave more like normal ocean lithosphere, where serpentinization is commonly confined to the uppermost mantle^{23,24}. We therefore speculate that differences in lithospheric composition favour serpentinization of the upper mantle in amagmatic lithosphere but limit serpentinization of magmatic lithosphere at the same depth levels. Alternatively, there may be differences in the connectivity of the fluid pathways that enable or prevent deeply penetrating water circulation in amagmatic and magmatic lithosphere, respectively.

To assess the relevance of our local seismicity surveys we examined the teleseismic earthquake record of ultraslow-spreading ridges⁵. The average numbers of earthquakes per kilometre of rift axis (for calculation see Methods) confirm the different deformation styles of predominantly magmatic and amagmatic rift segments (Figs 1 and 2 and Extended Data Fig. 4). Figure 2 and Extended Data Fig. 4 illustrate these differences for Gakkel Ridge and the SWIR, respectively: more abundant, stronger and often clustered earthquakes coincide with basalt exposure and a strong central magnetic anomaly in the magmatic Western Volcanic Zone, whereas reduced seismicity correlates with the occurrence of peridotite and an absence of magnetic anomalies in the Sparsely Magmatic Zone. Earthquakes there tend to be connected to minor basalt exposures. Volcanic centres may host extensive teleseismic earthquake swarms despite their locally thin elastic lithosphere (Fig. 1, site 3). The repeated teleseismic earthquake swarms between 1996 and 2003 at the Segment-8 volcano potentially mark a phase of magmatic activity⁵. The complete absence of seismicity underneath the volcano

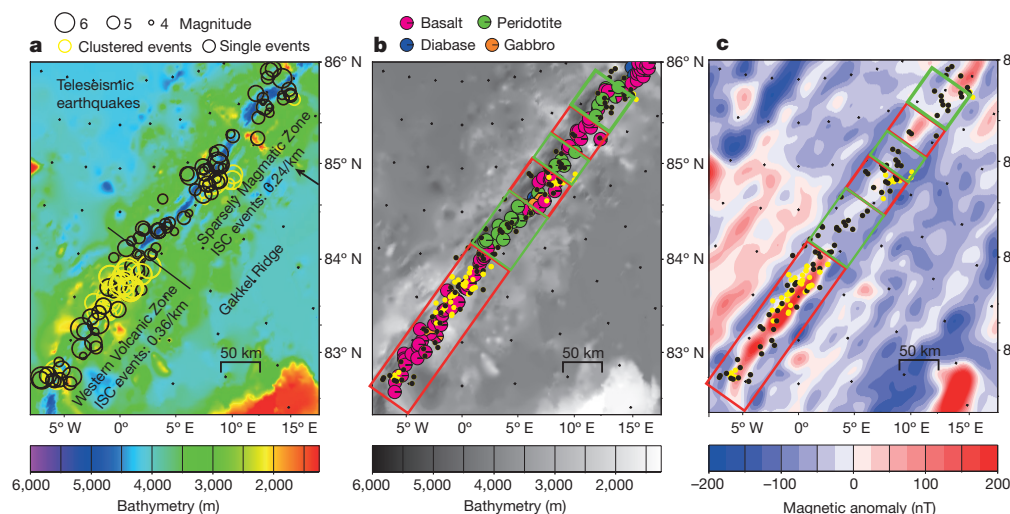


Figure 2 | Contrasting magmatic and amagmatic sections of Gakkel Ridge. **a**, Teleseismic earthquake activity (open circles, scaled with magnitude) over the bathymetry. Yellow circles mark earthquake clusters of two or more events related in time and space. Off-axis highs (arrow) extending away from the ridge axis indicate the long-term stability of this segmentation pattern. Teleseismic event rates are also indicated for comparison between sections. **b**, Dredge lithology. Data from ref. 4. **c**, Magnetic anomalies. Data from ref. 31. Earthquakes from **a** are shown by the dots. Magmatic lithosphere (red boxes) shows more and stronger earthquakes, basalt exposure and magnetic anomalies. Amagmatic lithosphere (green boxes) shows less seismicity, peridotite exposure and lacks a central magnetic anomaly.

during our survey in 2012–2013 may therefore be a result of increased temperatures caused by recent magmatism.

From the teleseismic and local seismicity records, we can thus define two end-member types of ocean lithosphere representative for ultraslow-spreading ridges (Figs 1–3). The first type, ‘amagmatic lithosphere’, has an elastic thickness of up to 35 km and lacks an igneous crust. Its serpentinized mantle rocks carry only weak magnetization²⁵ and are prone to deep-reaching serpentinization that results in aseismic deformation in the upper part of the lithosphere. Amagmatic lithosphere therefore shows a prominent reduction in seismicity. The second type, ‘magmatic lithosphere’, is thinner and shallows towards volcanoes. A thin, igneous crust carrying a remanent magnetization is present²⁵. Magmatic lithosphere is stronger, with considerable seismicity throughout its elastic portion. As it has experienced some melting and melt migration, we speculate that a different lithospheric composition may prevent extensive serpentinization. Dredge statistics from Gakkel Ridge and the SWIR show a higher percentage of gabbroic veins in magmatic lithosphere than in amagmatic lithosphere²⁶, potentially suggesting that more melts are being trapped at shallower levels in magmatic lithosphere (Fig. 3). However, as both extensive alteration to large depths and trapped melts could account for the observed low seismic velocities in our one-dimensional velocity models of the lithospheric mantle (Extended Data Figs 1b, 2b and 3b), our data can highlight only the different deformation styles of magmatic and amagmatic lithosphere, not determine their petrologic cause. Future high-resolution seismic studies combined with geological sampling are needed to determine the detailed velocity structure and composition of the end-member types of lithosphere described here.

However, the recognition of these lithosphere types and their geophysical characteristics leads to a conceptual advance in the understanding of ultraslow lithosphere accretion, which we sketch in Fig. 3 for Gakkel Ridge. The extensive circulation of water through amagmatic lithosphere may further cool and thicken the lithosphere locally, leading to a pronounced topography of its base that enables along-axis flow of melt on the segment scale towards the topographic shoals under magmatic zones. These magmatic zones vary in size from tiny patches in amagmatic zones (Fig. 1; site 1, 35 km and 90 km along the profile, see also the Sparsely Magmatic Zone in Figs 2 and 3), to volcanic centres with different thermal and magmatic states (Fig. 1, sites 2 and 3, and the Eastern Volcanic Zone in Fig. 3) and to robustly magmatic zones with extended along-axis magmatism (the Western Volcanic Zone in Figs 2 and 3). Lithospheric thinning is observed in all cases but the vertical amount and the lateral extent vary. The complete lack of seismicity at volcanic centres observed in our short survey period may indicate high temperatures connected to recent magmatic activity; it is likely to be transient given the observed teleseismic earthquake activity in these areas over longer periods. The variable appearance of magmatic lithosphere may therefore be caused by differences in lithospheric thickness, together with differences in the geometry and effectiveness of melt extraction²⁶, and by differences in melt availability, which is dependent on the mantle composition and fertility³. In addition to spatial variability, melt delivery is also expected to vary in time at a given location. As seismicity provides only snapshots in time of the lithospheric structure, the variable appearance of magmatic lithosphere also reflects different stages in its temporal evolution.

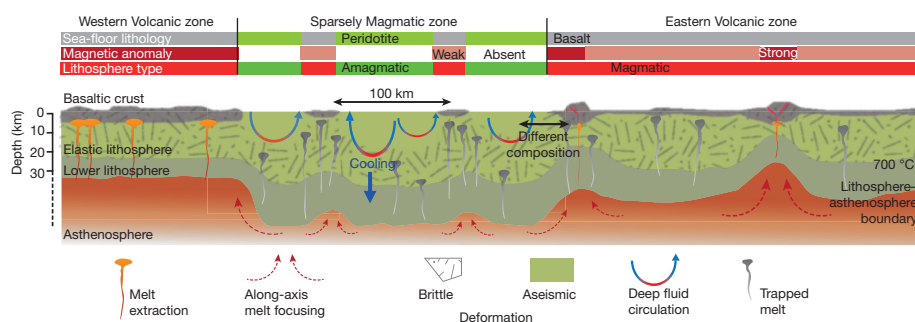


Figure 3 | Conceptual sketch of the two lithosphere types at the Gakkel Ridge. Scales are approximate. Deep-reaching alteration in amagmatic ridge sections cools and further thickens the lithosphere. The pronounced topography of the lithosphere–asthenosphere boundary focuses melts towards magmatic sections. Melt ascent there may result in a different lithospheric composition that is less prone to alteration and demonstrates brittle behaviour, whereas amagmatic lithosphere can deform aseismically.

We show that lithospheric thinning in magmatic sections is not the only factor that produces a topography of the lithosphere–asthenosphere boundary. The proposed deep-reaching serpentinization of amagmatic lithosphere implies extensive water circulation that in turn results in the cooling of the lithosphere from above. Spatially variable cooling of the lithosphere may therefore also contribute to enhancing the topography of the lithosphere–asthenosphere boundary. Once this topography is established, the process is self-sustaining. Gakkel Ridge, Knipovich Ridge and the western SWIR show off-axis highs that extend away from volcanic centres in the spreading direction and a spatially stable seafloor magnetic anomaly pattern (Fig. 2), both of which document long-term stability of the lithospheric accretion modes. Any reorganization of this pattern as observed on the eastern SWIR requires large instabilities such as a rapid cooling and thickening of the lithosphere at magmatic centres to prevent further melt pooling¹⁷.

Our study provides geophysical constraints on the along-axis thermal structure of ultraslow spreading ridges, against which a multitude of petrologic models of ultraslow spreading^{18,26,27} can be validated. A further insight is that ultraslow spreading ocean lithosphere exhibits a completely different deformation mode that we attribute to alteration of mantle rocks reaching 15 km depth.

This deformation style distinguishes ultraslow-spreading ridges from faster-spreading ridges: slow-spreading ridges show a dominant pattern of higher levels of seismicity at the colder, magma-poor segment ends¹³, in contrast to what we observe at the ultraslow-spreading ridges where less seismicity is associated with amagmatic regions. Exhumation of mantle lithosphere along major detachment faults produces substantial earthquake activity at the Mid-Atlantic Ridge^{13,28}. The seismicity delineates the detachment faults to depths of 7 km into the shallow upper mantle²⁹. At ultraslow-spreading ridges, detachment faulting is also thought to be a fundamental process in mantle rock exhumation³⁰ and the generation of smooth sea floor⁶, but our study suggests that deformation along such shear zones occurs entirely aseismically owing to deep-reaching serpentinization, explaining the contrasting seismicity pattern of slow- and ultraslow-spreading ridges.

Serpentinization in young oceanic lithosphere is generally believed to be limited to shallow depths of 4–6 km below detachment faults, but its extent is generally difficult to estimate from seismic velocities or rock samples gained from dredging or drilling²⁴. Our data suggest that in the extensive amagmatic regions of ultraslow-spreading ridges serpentinization and fluid circulation may reach far deeper into the mantle than previously assumed.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 22 October 2015; accepted 18 April 2016.

Published online 29 June 2016.

1. Bown, J. W. & White, R. S. Variation with spreading rate of oceanic crustal thickness and geochemistry. *Earth Planet. Sci. Lett.* **121**, 435–449 (1994).
2. Dick, H., Lin, J. & Schouten, H. An ultraslow-spreading class of ocean ridge. *Nature* **426**, 405–412 (2003).
3. Sauter, D. & Cannat, M. The ultraslow spreading Southwest Indian Ridge. *Geophys. Monogr. Ser.* **188**, 153–173 (2010).
4. Michael, P. J. *et al.* Magmatic and amagmatic seafloor generation at the ultraslow-spreading Gakkel ridge, Arctic Ocean. *Nature* **423**, 956–961 (2003).
5. Schlindwein, V. Teleseismic earthquake swarms at ultraslow spreading ridges: indicator for dyke intrusions? *Geophys. J. Int.* **190**, 442–456 (2012).
6. Cannat, M. *et al.* Modes of seafloor generation at a melt-poor ultraslow-spreading ridge. *Geology* **34**, 605–608 (2006).
7. White, R. S., McKenzie, D. & O'Nions, R. K. Oceanic crustal thickness from seismic measurements and rare earth element inversions. *J. Geophys. Res.* **97**, 19683–19715 (1992).
8. Anderson, D. L. Lithosphere, asthenosphere, and perisphere. *Rev. Geophys.* **33**, 125–149 (1995).

9. Cannat, M. How thick is the magmatic crust at slow spreading oceanic ridges? *J. Geophys. Res.* **101**, 2847–2857 (1996).
10. Edmonds, H. N. *et al.* Discovery of abundant hydrothermal venting on the ultraslow-spreading Gakkel Ridge in the Arctic Ocean. *Nature* **421**, 252–256 (2003).
11. Cannat, M. Emplacement of mantle rocks in the seafloor at mid-ocean ridges. *J. Geophys. Res.* **98**, 4163–4172 (1993).
12. Niu, X. *et al.* Along-axis variation in crustal thickness at the ultraslow spreading Southwest Indian Ridge (50°E) from a wide-angle seismic experiment. *Geochem. Geophys. Geosyst.* **16**, 468–485 (2015).
13. Escartin, J. *et al.* Central role of detachment faults in accretion of slow-spreading oceanic lithosphere. *Nature* **455**, 790–794 (2008).
14. Schlindwein, V., Demuth, A., Geissler, W. H. & Jokat, W. Seismic gap beneath Logachev Seamount: indicator for melt focusing at an ultraslow mid-ocean ridge? *Geophys. Res. Lett.* **40**, 1703–1707 (2013).
15. Schlindwein, V. The expedition of the research vessel “Polarstern” to the Antarctic in 2013 (ANT-XXIX/8). *Rep. Polar Marine Res.* **672**, 14–22 (2014).
16. Dusunur, D. *et al.* Seismological constraints on the thermal structure along the Lucky Strike segment (Mid-Atlantic Ridge) and interaction of tectonic and magmatic processes around the magma chamber. *Mar. Geophys. Res.* **30**, 105–120 (2009).
17. Cannat, M., Rommevaux-Jestin, C. & Fujimoto, H. Melt supply variations to a magma-poor ultra-slow spreading ridge (Southwest Indian Ridge 61° to 69°E). *Geochem. Geophys. Geosyst.* **4**, 9104 (2003).
18. Montési, L. G. J., Behn, M. D., Hebert, L. B., Lin, J. & Barry, J. L. Controls on melt migration and extraction at the ultraslow Southwest Indian Ridge 10°–16°E. *J. Geophys. Res.* **116**, B10102 (2011).
19. Standish, J. J., Dick, H. J. B., Michael, P. J., Nelson, W. G. & O'Hearn, T. MORB generation beneath the ultraslow spreading Southwest Indian Ridge (9–25°E): major element chemistry and the importance of process versus source. *Geochem. Geophys. Geosyst.* **9**, Q05004 (2008).
20. Cannat, M., Rommevaux-Jestin, C., Sauter, D., Deplus, C. & Mendel, V. Formation of the axial relief at the very slow spreading Southwest Indian Ridge (49° to 69°E). *J. Geophys. Res.* **104**, 22825–22843 (1999).
21. Amiguet, E. *et al.* Creep of phyllosilicates at the onset of plate tectonics. *Earth Planet. Sci. Lett.* **345–348**, 142–150 (2012).
22. Schwartz, S. *et al.* Pressure–temperature estimates of the lizardite/antigorite transition in high pressure serpentinites. *Lithos* **178**, 197–210 (2013).
23. Escartin, J., Hirth, G. & Evans, B. Strength of slightly serpentinized peridotites: implications for the tectonics of oceanic lithosphere. *Geology* **29**, 1023–1026 (2001).
24. Rouméjon, S. & Cannat, M. Serpentinization of mantle-derived peridotites at mid-ocean ridges: mesh texture development in the context of tectonic exhumation. *Geochem. Geophys. Geosyst.* **15**, 2354–2379 (2014).
25. Sauter, D., Cannat, M. & Mendel, V. Magnetization of 0–26.5 Ma seafloor at the ultraslow spreading Southwest Indian Ridge, 61°–67°E. *Geochem. Geophys. Geosyst.* **9**, Q04023 (2008).
26. Sleep, N. H. & Warren, J. M. Effect of latent heat of freezing on crustal generation at low spreading rates. *Geochem. Geophys. Geosyst.* **15**, 3161–3174 (2014).
27. Cannat, M. *et al.* Spreading rate, spreading obliquity, and melt supply at the ultraslow spreading Southwest Indian Ridge. *Geochem. Geophys. Geosyst.* **9**, Q04002 (2008).
28. Simão, N. *et al.* Regional seismicity of the Mid-Atlantic Ridge: observations from autonomous hydrophone arrays. *Geophys. J. Int.* **183**, 1559–1578 (2010).
29. deMartin, B. J., Sohn, R. A., Pablo Canales, J. & Humphris, S. E. Kinematics and geometry of active detachment faulting beneath the Trans-Atlantic Geotraverse (TAG) hydrothermal field on the Mid-Atlantic Ridge. *Geology* **35**, 711–714 (2007).
30. Sauter, D. *et al.* Continuous exhumation of mantle-derived rocks at the Southwest Indian Ridge for 11 million years. *Nat. Geosci.* **6**, 314–320 (2013).
31. Maus, S. *et al.* EMAG2: a 2-arc min resolution Earth Magnetic Anomaly Grid compiled from satellite, airborne, and marine magnetic measurements. *Geochem. Geophys. Geosyst.* **10**, Q08005 (2009).

Acknowledgements This study was enabled by grants SCHL853/1-1 and SCHL853/3-1 of the German Science Foundation to V.S. Instruments were borrowed from the DEPA pool. We acknowledge the efforts of the crews of RV *Polarstern* cruises ANT-XXIX/2+8 and ARK-XXIV/3, RV *Meteor* cruise M101 and RV *Marion Dufresne*.

Author Contributions V.S. planned and conducted the surveys, processed data for site 3 and wrote the paper. F.S. processed data from site 1. Both authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to V.S. (Vera.Schlindwein@awi.de).

Reviewer Information *Nature* thanks S. M. Carbotte and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Data processing. Seismic signals were identified with a short time average/long time average trigger in the continuous data stream of the OBSs, optimized for detecting all local earthquakes. All of the triggered events were reviewed by an analyst and spurious events were removed. P- and S-wave arrivals of all of the local earthquakes that were recorded by three or more stations were subsequently hand-picked and located with the linear least-squares algorithm Hyposat³², which also constrains the inversion with S-P travel-time differences.

One-dimensional velocity model. To derive a minimum one-dimensional velocity–depth profile, as commonly used in earthquake location³³, we used the results of refraction seismic surveys^{15,34,35} conducted within each of the study sites (Extended Data Figs 1a, 2a and 3a). We extracted smoothed average velocity–depth profiles for the upper 5 km, where the refraction seismic data provided sufficient ray cover. We thus constructed an initial velocity model and located all of the earthquakes. We used this preliminary location run to select a subset of well recorded events of at least 5 km depth that are situated within the network of stations. These events are considered the most sensitive to changes in the velocity model. We then tested a wide range of conceivable sub-Mohorovičić discontinuity (sub-Moho) velocities but kept the velocity model in the crust fixed as constrained by refraction seismic data. We located the subset of well constrained events for each velocity model and assessed the performance of each velocity model. We selected the velocity model that provided the lowest average root mean squared travel-time residual while still locating a large number of events. This final velocity model (Extended Data Figs 1b, 2b and 3b, red) was then used to locate all earthquakes.

Robustness tests. To assess the effects of the choice of velocity model on the location results and in particular to evaluate the reliability of the large hypocentre depths obtained, we performed several robustness tests. We located all of the events with a slow end-member velocity model with velocities reduced by 0.3 km s^{-1} compared with the final model and a fast end-member velocity model that consisted of a 4-km-thick crust underlain by velocities of 8 km s^{-1} , as is common for oceanic crust⁷ (Extended Data Figs 1b, 2b and 3b). Both low and high end-member velocity models produced a considerably poorer fit to the phase arrivals of all earthquakes as seen from the average root mean squared travel time residual in Extended Data Table 1. Fast sub-Moho velocities resulted in a failure to determine the hypocentre depth of many events at all three locations. The location algorithm had to fix the hypocentre depth to converge on the results. The number of well-determined hypocentre depths with depth errors of less than 5 km is therefore much smaller for the high-velocity model and illustrates its inappropriateness.

We further examined the effect of variations in the velocity model by $\pm 0.3 \text{ km s}^{-1}$ on hypocentre depth. The reduced velocity model is identical to the low-velocity end-member model tested above (Extended Data Figs 1b, 2b and 3b, orange). For the faster-velocity model we increased the velocities of the final velocity model by $+0.3 \text{ km s}^{-1}$ throughout but kept 8.0 km s^{-1} at 40 km depth fixed (Extended Data Figs 1b, 2b and 3b, purple) to avoid the convergence problems described above. Faster velocities again produced a poorer fit to the data and yielded fewer reliable hypocentre depths (Extended Data Table 2). These, however, are on average slightly shallower than hypocentres located with lower-velocity models. Changes in the average hypocentre depth are in all cases smaller than the average depth error and they are much smaller than the along-axis depth variations of the band of seismicity shown in Fig. 1 and interpreted here. Prominent deviations from the one-dimensional minimum velocity model due to local heterogeneities in the subsurface result in non-zero average station residuals rather than substantially shifting hypocentres. On site 3, the station located on the crest of the volcano showed a positive S-phase residual of on average 0.232 s. Omitting this station from the location procedure did not distort the pattern of hypocentres interpreted.

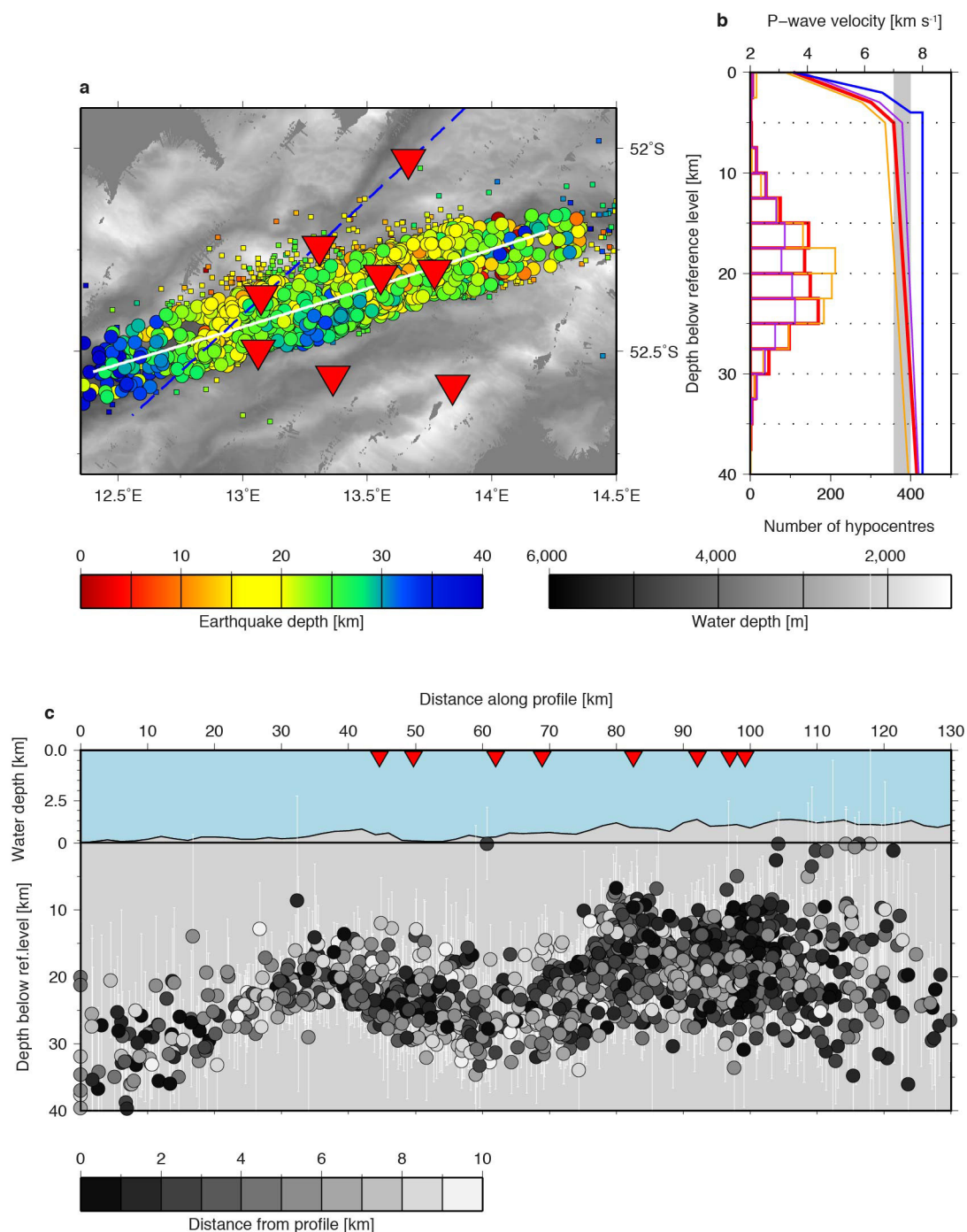
Selection and display of hypocentres. Extended Data Figs 1–3 show all 5,379 located earthquakes with a hypocentre solution, irrespective of location accuracy. These figures hence include also all smaller events. For Fig. 1 we imposed a maximum hypocentre depth error of 5 km as the sole quality criterion, fulfilled by 903, 441 and 2,625 events in data sets 1–3, respectively. The cross-sections displayed in Fig. 1 are based on a total of 3,664 events within 9.4 km of the cross-section profile. These events have a mean horizontal error of $\pm 3.5 \text{ km}$, a mean depth error of $\pm 2.7 \text{ km}$ and an average root mean squared travel-time residual of 0.24 s. Their locations were calculated from on average 15.5 phases at 6.3 stations, the closest station being on average 7.2 km away. For the majority of the events, the distance to the closest recording station is hence smaller than the hypocentre depth, which is generally considered necessary to obtain good hypocentre solutions. In Fig. 1, we highlight events outside the network where the distance to the next station becomes larger than the hypocentre depth. For these events, hypocentre depths have to be interpreted with care.

The band of seismicity was in all cases more or less flat-lying in the across-axis direction. Therefore, the projection of the hypocentres onto the cross-sections did not distort the seismicity pattern. Extended Data Figs 1c, 2c and 3c reveal no systematic dependence of hypocentre depth on projection distance. Aseismic regions interpreted here are furthermore not a consequence of earthquake selection as they appear in the full data set (Extended Data Figs 1–3) in the same way as in Fig. 1, indicating that these areas are also devoid of the weaker events that were omitted from Fig. 1.

Teleseismic events. Teleseismic earthquakes are used from a compilation⁵ of reviewed locations of the Bulletin of the International Seismological Centre occurring from 1976 to 2010 within 30 km of the rift axis for the Gakkel and Knipovich ridges and within 35 km of the rift axis for the SWIR owing to the larger location uncertainties there. A single-link cluster analysis identified earthquake clusters in time and space. Owing to their long observation period, teleseismic earthquake catalogues of mid-ocean ridges can reproduce the main features of along-axis seismicity variations on a regional scale that are visible in more complete catalogues of hydroacoustically recorded events²⁸. We project the event locations onto our profiles in Fig. 1 and count the number of teleseismic events in our 35-year-long catalogue in bins of 13 km, displayed as histograms in Fig. 1. We further calculate the average number of events per kilometre of the rift axis for comparison of the considered rift sections in Figs 1 and 2 and Extended Data Fig. 4. As the mean location uncertainty of the teleseismic events is about $\pm 31 \text{ km}$ for the SWIR compared with $\pm 12 \text{ km}$ at Gakkel Ridge⁹, we cannot unambiguously assign individual earthquakes to either amagmatic or magmatic subsections of the oblique supersegment, in particular. We therefore calculate seismicity rates for entire segments (Fig. 2 and Extended Data Fig. 4) or major portions thereof (Fig. 1) for comparison with adjacent segments. The average seismicity rates of the predominantly amagmatic supersegments contain some earthquakes that are connected with minor occurrences of magmatic lithosphere there. Purely amagmatic lithosphere may have even lower seismicity rates and the difference in the seismicity rate between amagmatic and magmatic lithosphere may be more pronounced.

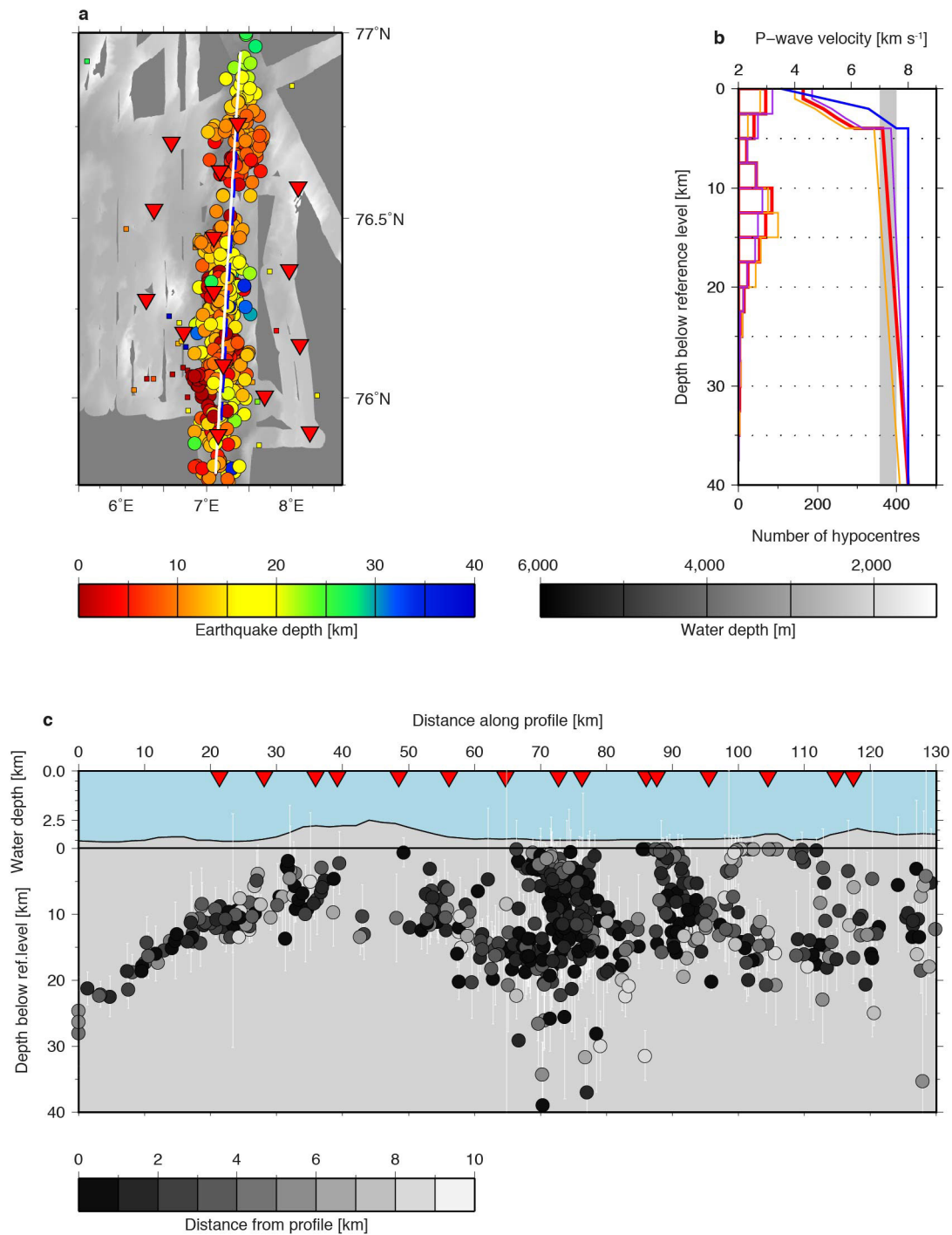
Figure 2 and Extended Data Fig. 4 show the individual teleseismic earthquake locations compared with a global map of magnetic anomalies³¹. Events that are part of clusters of two or more events occurring in close relation in time or space are highlighted. Note that such clusters occur only in regions of magmatic lithosphere. **Estimates of lithosphere temperature.** Absolute temperatures at the transition depth from brittle to ductile behaviour are discussed by several authors^{8,36,37} and their estimates range between 550°C and 750°C . For mid-ocean-ridge lithosphere compositions, preference is given to higher temperatures within this range⁹. As mantle hotter than about 650°C cannot build up long-term stresses, the base of seismicity and the transition between ductile and brittle rheologies have been associated with isotherms⁸. We therefore assume here that the maximum depth of seismicity delineates an isotherm of about 700°C . This value was chosen arbitrarily, the exact temperature is irrelevant for our conclusions. To obtain a crude estimate of the depth of the 400°C isotherm in limited along-axis areas we assume a constant temperature gradient between the sea floor and the depth of the 700°C isotherm. We furthermore assume that the $1,200^\circ\text{C}$ isotherm, used as a proxy for the lithosphere–asthenosphere boundary, has roughly the same along-axis topography as the 700°C isotherm. Temperature fields calculated for mid-ocean-ridge axes^{26,38,39} show approximately constant spacing of isotherms in the brittle lithosphere and similar shapes of the 700°C and $1,200^\circ\text{C}$ isotherms, justifying our assumptions.

32. Schweitzer, J. HYPOSAT—an enhanced routine to locate seismic events. *Pure Appl. Geophys.* **158**, 277–289 (2001).
33. Kissling, E., Ellsworth, W. L., Eberhart-Phillips, D. & Kradolfer, U. Initial reference models in local earthquake tomography. *J. Geophys. Res.* **99**, 19635–19646 (1994).
34. Jokat, W., Kollofrath, J., Geissler, W. H. & Jensen, L. Crustal thickness and earthquake distribution south of the Logachev Seamount, Knipovich Ridge. *Geophys. Res. Lett.* **39**, L08302 (2012).
35. Minshall, T. A., Muller, M. R. & White, R. S. Crustal structure of the Southwest Indian Ridge at 66°E : seismic constraints. *Geophys. J. Int.* **166**, 135–147 (2006).
36. McKenzie, D., Jackson, J. & Priestley, K. Thermal structure of oceanic and continental lithosphere. *Earth Planet. Sci. Lett.* **233**, 337–349 (2005).
37. Chen, W.-P. & Molnar, P. Focal depths of intracontinental and intraplate earthquakes and their implications for the thermal and mechanical properties of the lithosphere. *J. Geophys. Res.* **88**, 4183–4214 (1983).
38. Chen, W.-P. & Molnar, P. A nonlinear rheology model for mid-ocean ridge axis topography. *J. Geophys. Res.* **95**, 17583–17604 (1990).
39. Montési, L. G. J. & Behn, M. D. Mantle flow and melting underneath oblique and ultraslow mid-ocean ridges. *Geophys. Res. Lett.* **34**, L24307 (2007).



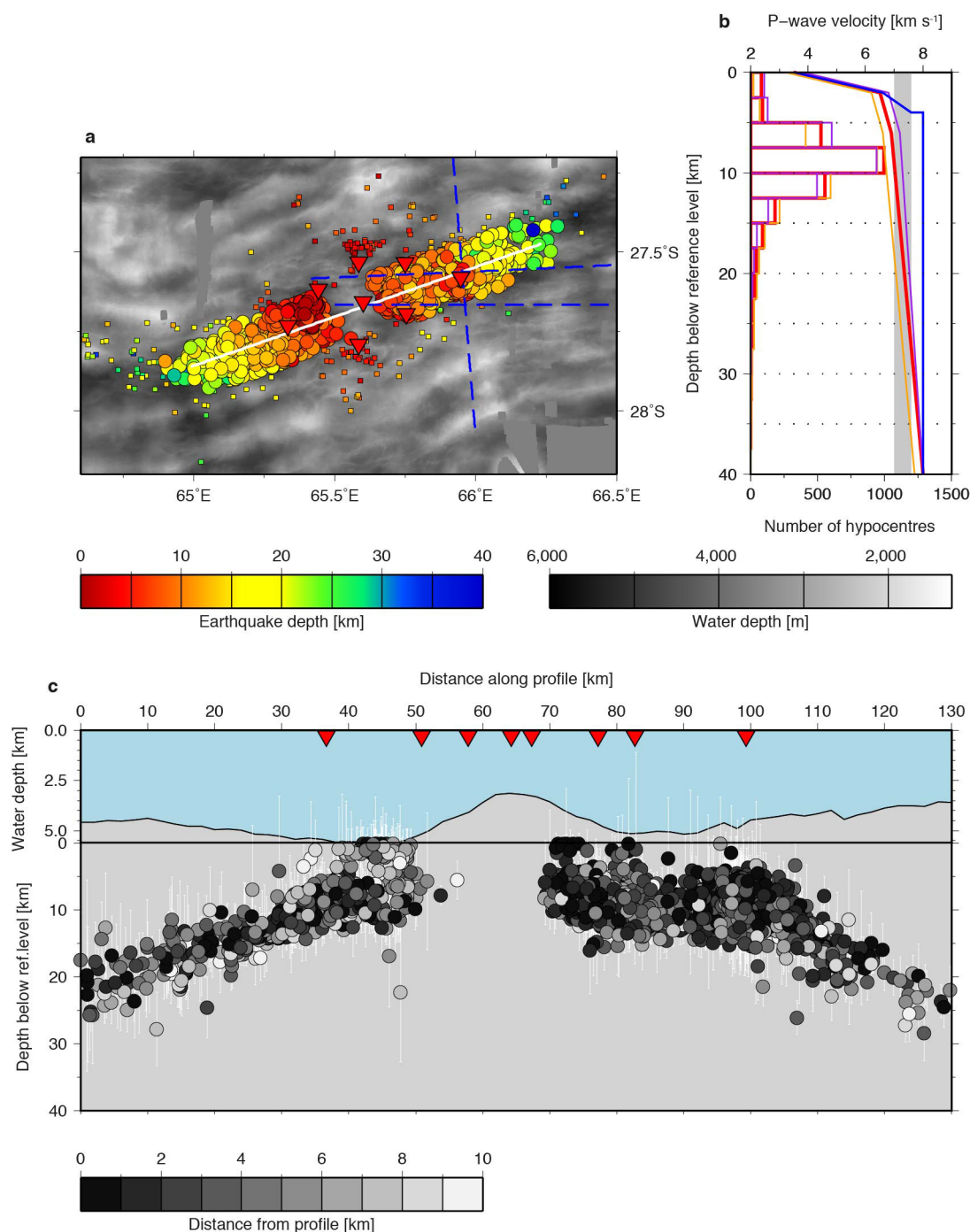
Extended Data Figure 1 | Earthquake location at survey site 1 including poorly located events. **a**, Epicentres (circles) colour-coded by hypocentre depth. Earthquakes not projected onto the cross-section (white line) in **c** are shown by squares. The red inverted triangles show OBS locations and the dashed blue line shows the position of the refraction seismic line¹⁵ used to constrain velocities in the uppermost lithosphere. **b**, The final velocity model used for earthquake location is shown in red, and the velocity models used for the robustness tests are shown in blue (the fast end-member representing a thin crust with ultramafic rocks underneath) orange (a velocity reduction of 0.3 km s⁻¹ relative to the final model) and purple (a velocity increase of 0.3 km s⁻¹ relative to the final model).

Velocities of 7.0–7.6 km s⁻¹ (grey bar) are considered anomalously low for lithospheric mantle. The histograms show the distribution of hypocentre depths obtained for the different velocity models. Faster models result in fewer well-located events, but the depth distribution is similar (see Extended Data Table 2). **c**, Cross-section of the hypocentres projected onto the axis and colour-coded according to the distance from the profile. The topography of the seismicity band is not an artefact of projection because at all depth intervals the earthquakes from various off-axis distances are present. The aseismic regions remain devoid of seismicity even when all poorly located earthquakes are shown.



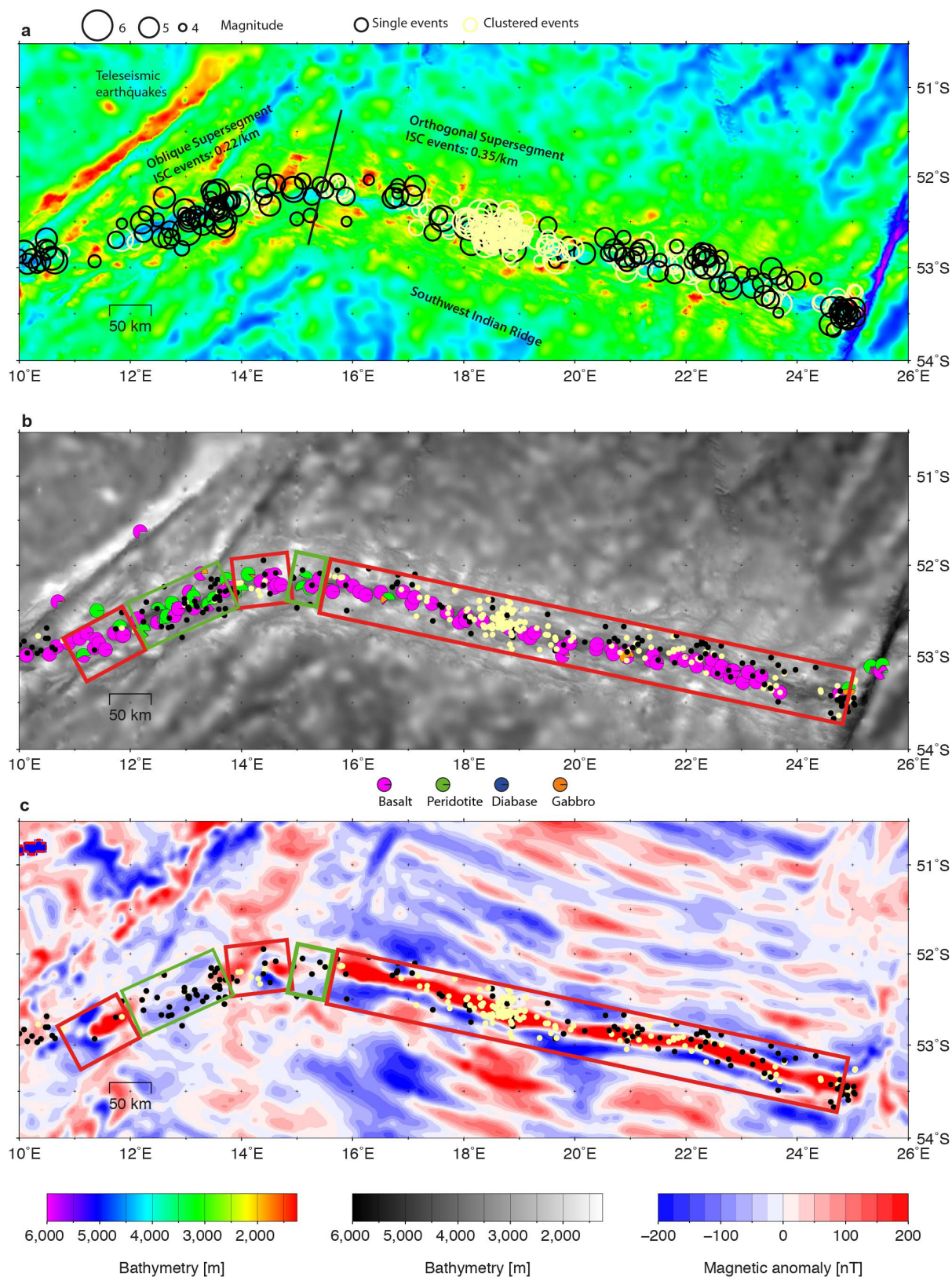
Extended Data Figure 2 | Earthquake location at survey site 2 including poorly located events. **a**, Epicentres (circles) colour-coded by hypocentre depth. Earthquakes not projected onto the cross-section (white line) in **c** are shown by the squares. The red inverted triangles show OBS locations and the dashed blue line indicates the position of the refraction seismic line³⁴ used to constrain velocities in the uppermost lithosphere. **b**, The final velocity model used for the earthquake location is shown in red and the velocity models used for the robustness tests are shown in blue (the fast end-member representing a thin crust with ultramafic rocks underneath) orange (a velocity reduction of 0.3 km s^{-1} relative to the final model) and purple (a velocity increase of 0.3 km s^{-1} relative to the final model).

Velocities of $7.0\text{--}7.6 \text{ km s}^{-1}$ (grey bar) are considered as anomalously low for lithospheric mantle. The histograms show the distribution of hypocentre depths obtained for the different velocity models. Faster models result in fewer well located events, but the depth distribution is similar (see Extended Data Table 2). **c**, Cross-section of the hypocentres projected onto the axis and colour-coded according to the distance from the profile. The topography of the seismicity band is not an artefact of projection because for all depth intervals the earthquakes from various off-axis distances are present. The aseismic regions remain devoid of seismicity even when all poorly located earthquakes are shown.



Extended Data Figure 3 | Earthquake location at survey site 3 including poorly located events. **a**, Epicentres (circles) colour-coded by hypocentre depth. Earthquakes not projected onto the cross-section (white line) in **c** are shown by the squares. The red inverted triangles show OBS locations and the dashed blue line indicates the position of the refraction seismic lines³⁵ used to constrain velocities in the uppermost lithosphere. **b**, The final velocity model used for the earthquake location is shown in red and the velocity models used for the robustness tests are shown in blue (the fast end-member representing a thin crust with ultramafic rocks underneath) orange (a velocity reduction of 0.3 km s^{-1} relative to the final model) and purple (a velocity increase of 0.3 km s^{-1} relative to the final model).

Velocities of $7.0\text{--}7.6 \text{ km s}^{-1}$ (grey bar) are considered as anomalously low for lithospheric mantle. The histograms show the distribution of hypocentre depths obtained for the different velocity models. Faster models result in fewer well-located events, but the depth distribution is similar (see Extended Data Table 2). **c**, Cross-section of the hypocentres projected onto the axis and colour-coded according to the distance from the profile. The topography of the seismicity band is not an artefact of projection because at all depth intervals the earthquakes from various off-axis distances are present. The aseismic regions remain devoid of seismicity even when all poorly located earthquakes are shown.



Extended Data Figure 4 | Contrasting magmatic and amagmatic sections of western SWIR. **a**, Teleseismic earthquake activity (open circles, scaled with magnitude) over bathymetry. The light yellow circles mark earthquake clusters of two or more events that are related in time and space. Data from ref. 5. **b**, Dredge lithology. Data from ref. 2. **c**, Magnetic anomalies. Data from ref. 31. Earthquakes from **a** are shown by the dots. The predominantly magmatic orthogonal supersegment shows

more abundant and often clustered teleseismic earthquakes and a marked magnetic anomaly. The predominantly amagmatic oblique supersegment shows less seismicity and peridotite exposure. Areas of magmatic and amagmatic lithosphere within this segment are defined from the seafloor lithology and magnetic patterns. The differences in the event rates within segments (see Fig. 2) are not visible here owing to a large uncertainty in earthquake locations.

Extended Data Table 1 | Location performance of the final one-dimensional velocity models and end-member models

Site Velocity model	Number of earthquakes located	Average rms residual [s]	Number of earthquakes with depth solution	Number of earthquakes with depth error < 5 km
Site 1				
Slow-end	1960	0.502	1896	990
Final	1959	0.485	1871	903
Fast-end	1959	0.546	1556	242
Site 2				
Slow-end	737	0.286	631	461
Final	736	0.264	614	441
Fast-end	744	0.364	441	181
Site 3				
Slow-end	2974	0.226	2880	2557
Final	2974	0.196	2894	2625
Fast-end	2970	0.235	1998	1555

The final velocity model of each site achieved the best fit to all of the observed phases. The high end-member velocity models produce fewer and unstable hypocentre depths.

Extended Data Table 2 | Dependence of the hypocentre depths on the velocity model

Site Velocity model	Number of earthquakes located	Number of earthquakes with depth error < 5 km	Average depth [km]	Average depth error [km]
Site 1				
- 0.3 km/s	1960	990	20.3	3.8
Final	1959	903	20.2	4.0
+ 0.3 km/s	1960	629	20.0	4.1
Site 2				
- 0.3 km/s	737	461	12.2	2.8
Final	736	441	10.9	2.6
+ 0.3 km/s	738	403	9.6	2.7
Site 3				
- 0.3 km/s	2974	2557	10.6	2.5
Final	2974	2625	9.5	2.3
+ 0.3 km/s	2970	2497	8.8	2.3

Only inverted hypocentre depths with a depth error of less than 5 km are included in the average values. The individual distributions are shown in Extended Data Figs 1b, 2b and 3b for each site. Note that the changes in average depth due to variations of the velocity model are in all cases smaller than the average depth error.

Molecular logic behind the three-way stochastic choices that expand butterfly colour vision

Michael Perry¹, Michiyo Kinoshita², Giuseppe Saldi³, Lucy Huo¹, Kentaro Arikawa² & Claude Desplan^{1,3}

Butterflies rely extensively on colour vision to adapt to the natural world. Most species express a broad range of colour-sensitive Rhodopsin proteins in three types of ommatidia (unit eyes), which are distributed stochastically across the retina^{1–3}. The retinas of *Drosophila melanogaster* use just two main types, in which fate is controlled by the binary stochastic decision to express the transcription factor Spineless in R7 photoreceptors⁴. We investigated how butterflies instead generate three stochastically distributed ommatidial types, resulting in a more diverse retinal mosaic that provides the basis for additional colour comparisons and an expanded range of colour vision. We show that the Japanese yellow swallowtail (*Papilio xuthus*, Papilionidae) and the painted lady (*Vanessa cardui*, Nymphalidae) butterflies have a second R7-like photoreceptor in each ommatidium. Independent stochastic expression of Spineless in each R7-like cell results in expression of a blue-sensitive (Spineless^{ON}) or an ultraviolet (UV)-sensitive (Spineless^{OFF}) Rhodopsin. In *P. xuthus* these choices of blue/blue, blue/UV or UV/UV sensitivity in the two R7 cells are coordinated with expression of additional Rhodopsin proteins in the remaining photoreceptors, and together define the three types of ommatidia. Knocking out *spineless* using CRISPR/Cas9 (refs 5, 6) leads to the loss of the blue-sensitive fate in R7-like cells and transforms retinas into homogeneous fields of UV/UV-type ommatidia, with corresponding changes in other coordinated features of ommatidial type. Hence, the three possible outcomes of Spineless expression define the three ommatidial types in butterflies. This developmental strategy allowed the deployment of an additional red-sensitive Rhodopsin in *P. xuthus*, allowing for the evolution of expanded colour vision with a greater variety of receptors^{7,8}. This surprisingly simple mechanism that makes use of two binary stochastic decisions coupled with local coordination may prove to be a general means of generating an increased diversity of developmental outcomes.

Papilio butterflies have some of the most complex retinal mosaics of any insect^{1,2}. In *P. xuthus*, three ommatidial types express five Rhodopsins used in colour vision, motion detection, and in polarized light vision¹ (Fig. 1a, b). In contrast, *V. cardui* exhibits a somewhat less complex pattern of expression of just three Rhodopsins². The adult butterfly retina is tiered with four distal photoreceptors (butterfly R1–R4, bR1–4) just below the lens, four proximal photoreceptors bR5–R8 deeper in the ommatidium, and the bR9 cell below¹ (Fig. 1a).

The three stochastically distributed ommatidial types found in butterflies are reminiscent of the simpler *D. melanogaster* eye, which contains just two stochastic types⁹. In *D. melanogaster*, the stochastic variations are restricted to the ‘inner’ photoreceptors R7 and R8, in ommatidia in which R7 expresses UV-sensitive Rhodopsin Rh3, R8 expresses blue-sensitive Rhodopsin Rh5. When UV-sensitive Rhodopsin Rh4 is expressed in R7, R8 expresses green-sensitive Rhodopsin Rh6 (refs 10, 11). The remaining ‘outer’ photoreceptors R1–6 all express the same broad spectrum Rhodopsin Rh1 in all ommatidia and are involved in motion vision¹⁰. The stochastic cell

fate decision that determines ommatidial type in *D. melanogaster* is controlled by the transcription factor Spineless (Ss), and the stochastic expression of this transcription factor in R7 directly regulates Rh4 (refs 4, 12). This binary choice in R7 is then communicated to the underlying R8, thereby defining the two types of ommatidia^{11,13}.

Butterflies rely on colour vision for recognizing flowers, finding food, identifying mates, and in locating oviposition sites—challenging visual environments which are reflected by the complexity of their retinas. The increased ommatidial diversity in *P. xuthus* may have been useful for deploying a newly evolved red Rhodopsin, which is

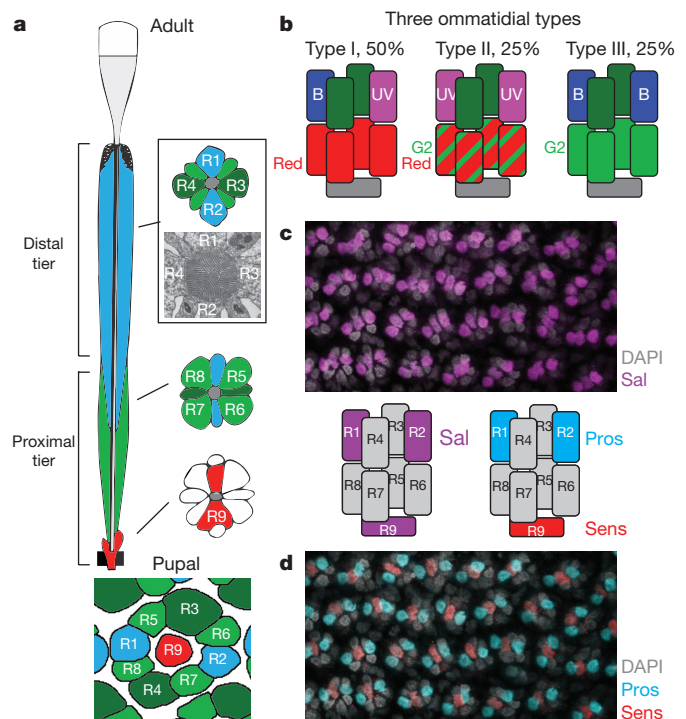


Figure 1 | Rhodopsin expression and ommatidial types in the *P. xuthus* retina. **a**, Cell bodies of the distal tier (bR1–4) and proximal tier (bR5–8 plus bR9) photoreceptors surround a central fused rhabdom (grey; transmission electron microscopy image shown in inset). **b**, Adult Rhodopsin expression in the three stochastically distributed ommatidial types. Combinations of UV-sensitive (UV) and blue-sensitive (B) Rhodopsins expressed in the distal tier are coordinated with specific long wavelength green-sensitive and red-sensitive Rhodopsins in the proximal tier. **c**, Antibodies against Sal (magenta) in pupal eye discs label three photoreceptor nuclei per ommatidium, including the central bR9. **d**, Sens (red) and Pros (cyan) expression provide evidence that there are two dR7-like cells in *P. xuthus* (bR1–2) and that bR9 is homologous to dR8: the additional ninth photoreceptor is a second dR7-like photoreceptor.

¹Department of Biology, New York University, New York, New York 10003, USA. ²Laboratory of Neuroethology, Department of Evolutionary Studies of Biosystems, SOKENDAI, The Graduate University for Advanced Studies, Hayama, Kanagawa 240-0115, Japan. ³New York University Abu Dhabi, Saadiyat Island, Abu Dhabi, United Arab Emirates.

expressed differently in each of the three ommatidial types^{7,14}. We set out to investigate how butterflies produce more than two stochastically distributed ommatidial types.

We sequenced the genomes and adult head transcriptomes of *P. xuthus* and *V. cardui*. *P. xuthus* has an established history being used for vision research^{1,15}, while *V. cardui* is a cosmopolitan species that can be maintained on artificial diet in the laboratory. Our results focus primarily on *P. xuthus*, with corresponding results from *V. cardui* shown in Extended Data Fig. 1. We produced antibodies against factors that define photoreceptor types in *D. melanogaster*, seeking to identify homologous photoreceptors in butterflies (Fig. 1)¹⁶.

The organization of butterfly photoreceptors during pupal stages resembles that of *D. melanogaster*, except for the presence of nine photoreceptors instead of eight (Fig. 1a). In *D. melanogaster*, the transcription factor Spalt is expressed specifically in colour-sensitive R7 and R8 (dR7 + 8) and is required for specifying these cells, which have long visual fibres (LVF) projecting to the medulla neuropil¹⁷. Outer photoreceptors dR1–6 have short visual fibres (SVF) terminating in the lamina neuropil¹⁷. A cross-reactive antibody to *D. melanogaster* Spalt labels three photoreceptor nuclei per ommatidium in pupal *P. xuthus* retinas (Fig. 1c and Extended Data Fig. 1c). Two of these photoreceptors become part of the distal tier of the adult ommatidium and correspond to the butterfly photoreceptors named R1 and R2 (bR1 + 2), while the central labelled nucleus is R9 (bR9). This expression pattern fits with previous data that suggested bR1 + 2 (and possibly bR9) are LVF photoreceptors¹⁸.

We identified the *P. xuthus* homologue of transcription factors Senseless (Sens), which is specifically expressed in *D. melanogaster* R8 (ref. 19), and Prospero (Pros), which marks R7 in flies²⁰. Antibodies against *P. xuthus* Sens specifically labelled bR9, indicating that this cell is homologous to dR8 (Fig. 1c). Antibodies against *P. xuthus* Pros marked two photoreceptors in butterflies, bR1 + 2 (Fig. 1c and Extended Data Fig. 1a), which are thus homologous to two dR7 photoreceptors, as had been proposed previously based on morphology¹⁶. The presence of Sal, Pros, and Sens in developing butterfly retinas suggests that the regulatory code that defines the different photoreceptor types is deeply conserved, with the additional ninth photoreceptor of butterflies being a second dR7-type.

The three *P. xuthus* ommatidial types (Fig. 1b)^{2,3} express different wavelength-specific Rhodopsin proteins not only in inner, but also in outer photoreceptors, suggesting that both LVF and SVF photoreceptors are involved in colour vision^{8,18}. In all three types of *P. xuthus* ommatidia, the two distal SVF photoreceptors bR3/R4 coexpress two green Rhodopsins, *P. xuthus* G1 (PxG1) and PxG2 (refs 1, 14). The Rhodopsins expressed in bR9, a cell that appears to play a reduced role in butterfly vision with little contribution to the rhabdom, have not been identified. In type I ommatidia (~50%), one of the dR7-like photoreceptors expresses PxUV and the other PxBlue, while the four proximal SVF photoreceptors bR5–8 express red PxR. In type II ommatidia (~25%), both R7-like photoreceptors express PxUV, whereas bR5–8 express both PxRed and low levels of PxG2 (ref. 21). Type III (~25%) express PxBlue in both R7-like photoreceptors, and high levels of PxG2 in bR5–8 (Fig. 1b). The choices in bR5–8 SVF photoreceptors are therefore precisely coupled with the choices in the two dR7-like cells (Fig. 1b). In *D. melanogaster*, a stochastic choice to express UV-sensitive Rh3 or UV-sensitive Rh4 is made in R7, and this binary choice is communicated to the underlying R8 to express blue-sensitive or green-sensitive Rhodopsin, respectively¹¹. All outer photoreceptors uniformly express Rh1.

The presence of two dR7-like photoreceptors suggested a model for stochastic patterning that could produce three butterfly ommatidial types. In the *D. melanogaster* retina (Fig. 2a), an independent, cell-intrinsic choice to express Spineless yields two possible fates in R7, Spineless^{ON} or Spineless^{OFF} (Fig. 2f)⁴. Spineless directly activates Rh4, as well as the transcription factor Defective proventriculus (Dve), which represses Rh3 (ref. 22). While Dve is also expressed in all R1–6,

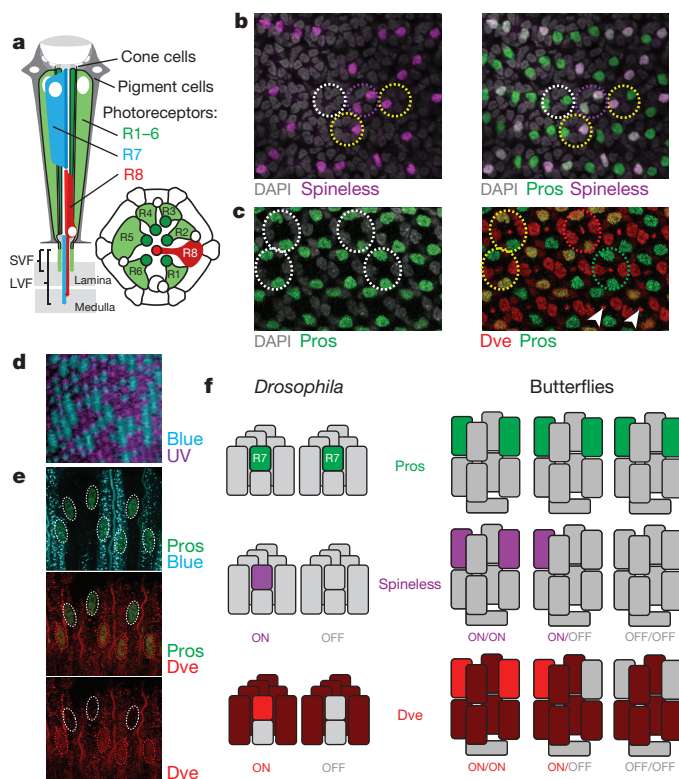


Figure 2 | Stochastic Spineless expression in flies and butterflies.

a, Schematic of an adult *D. melanogaster* ommatidium showing photoreceptor numbering and stacked dR7/R8. **b**, In *P. xuthus* pupal retinas, Spineless (magenta) is expressed stochastically in a subset of Pros-positive dR7-like cells (green). **c**, Dve (red) is expressed stochastically in a subset of dR7-like cells (labelled by Pros in green). Dve signal also localized to rhabdomeres (arrowheads), which is also observed in *D. melanogaster*¹². **d**, PxB and PxUV Rhodopsins are expressed stochastically in adult *P. xuthus* dR7-like cells, leading to a mosaic of ommatidia expressing B/B, B/UV or UV/UV, seen here in cross-section. **e**, Dve expression (red) in this lateral view of adult photoreceptors can be on or off in dR7-like nuclei (Pros in green). Dve is co-expressed with blue Rhodopsin PxBlue (cyan). The Dve signal again localized to the rhabdomeres as in Fig. 2c. **f**, In *D. melanogaster* (left), Spineless can be either on (purple) or off (grey) in *D. melanogaster* R7 cells (cyan). Dve is expressed in all SVF outer photoreceptors (dark red) and at a lower level in R7 cells (light red) that express Spineless. In butterflies (right), there are two dR7-like photoreceptors (green) per ommatidium in *P. xuthus*, as shown by Pros expression. Independent stochastic Spineless expression in these cells could yield three outcomes. Dve follows Spineless expression in each of the two dR7-like cells per ommatidium in *P. xuthus*. Schematics shown in **a** were adapted with permission from ref. 30.

it is on in Spineless^{ON} R7 cells and it is off in Spineless^{OFF} R7 (ref 12). If each butterfly R7-like cell were to also make an independent choice for Spineless expression, one would expect ON/ON, ON/OFF or OFF/OFF ommatidia, with Spineless matching Dve and Rhodopsin expression (Fig. 2f).

To test this possibility, we generated antibodies against *P. xuthus* Spineless and Dve, as well as against PxUV and PxBlue Rhodopsins. These antibodies cross-react with *V. cardui* proteins (see Extended Data Fig. 1). The Rhodopsin antibody stains showed the expected stochastic mosaic of three ommatidial types (Fig. 2d). Spineless expression in pupal retinas was also stochastic in the two Pros-expressing dR7-like photoreceptors, bR1 and bR2 (Fig. 2b). Spineless was either on in both bR1 and bR2 (ON/ON), ON/OFF, OFF/ON, or OFF/OFF. Dve was stochastically expressed in these dR7-like cells in an identical pattern (Figs 2c and 3c), suggesting that, as in flies, it is downstream of Spineless. We then attempted to determine which ommatidial types express Spineless and Dve in dR7-like cells. Although Spineless

expression was not detected in the adult when Rhodopsins were expressed, Dve was maintained specifically in cells that express PxBlue, but not in PxUV cells (Fig. 2e). Therefore, Spineless, Dve, and PxBlue are expressed in the same photoreceptors. This supports a model in which stochastic expression of Spineless in dR7-like cells defines the three ommatidial types: ON/OFF (and OFF/OFF) correspond to type I, OFF/OFF to type II and ON/ON to type III (Fig. 2f).

To test this model, we sought to knockout *spineless* expression using CRISPR/Cas9 (refs 5, 6). First, we determined the efficiency of the system by targeting *yellow*, a gene involved in melanin production. The bi-allelic knockout of *yellow* produced a strong wing and body yellow colour phenotype (Fig. 3a, b). Varying the timing of injection and concentration of single guide RNAs (sgRNAs) and Cas9 protein produced a range of effects, from the loss of almost all black-pigmented regions, to mixtures of mutant and non-mutant tissue (Fig. 3a, b). This type of mosaic knockout is useful for testing lethal mutations like *spineless*. We then targeted the *spineless* locus in both *P. xuthus* and *V. cardui*. The *spineless* knockout animals showed defects in antenna and were missing leg bristles (Extended Data Fig. 2), similar to *D. melanogaster* *spineless* mutant phenotypes²³. Sequencing revealed mutations at the target site (Extended Data Fig. 2). Unexpectedly, *spineless* also appeared to play a role in wing colour patterning. Colours produced via both melanin and ommochrome pathways were abolished in *spineless* mutant wing scales in both *P. xuthus* and *V. cardui* (Extended Data Fig. 2). Defects in the antennae and wings unfortunately meant that these animals did not eclose well and were not suitable for behavioural experiments.

In the retina, two independently tested sgRNAs targeting *spineless* caused loss of Spineless and Dve staining in dR7-like cells marked by Pros (Fig. 3c and Extended Data Fig. 1). Loss of Dve expression in dR7-like cells confirmed that Spineless acts upstream of Dve (Fig. 3c). Given the co-expression of Spineless, Dve, and PxBlue (Fig. 2e), we predicted the presence of uniform UV/UV type II ommatidia in *spineless* mutant tissues.

Wild-type *P. xuthus* retinas showed a mosaic of PxUV and PxBlue expression in the two dR7-like photoreceptors (Fig. 4a)²⁴. In contrast, large regions in *spineless* knockout retinas contained only PxUV/PxUV type ommatidia with complete loss of PxBlue expression (Fig. 4b). We then characterized the effect of *spineless* knockout on other aspects of ommatidial type such as the presence of non-visual filtering pigments, UV-induced fluorescence, and the expression of PxG1, PxG2, and PxR in SVF photoreceptors.

In wild-type *P. xuthus* retinas, either a red (in type I and II) or a yellow (in type III) pigment used for spectral tuning was expressed as a mosaic in the bR3–8 SVF photoreceptors (Fig. 4c, d)²⁵. In contrast, large regions in *spineless* knockout animals contained only red pigment-expressing ommatidia, in agreement with the loss of type III ommatidia (Fig. 4d).

UV light induces blue fluorescence in type II ommatidia due to the presence of fluorescing 3-hydroxyretinol in the wild-type retina²⁶ (Fig. 4e). The *spineless* knockout retinas contained large fields of ommatidia that all fluoresced blue (Fig. 4e), consistent with the loss of type I and III ommatidia.

The bR5–8 photoreceptors express either PxRed (type I), PxRed and low levels of PxG2 (type II), or high levels of PxG2 (type III)²¹ (Fig. 4c). In wild-type retinas, our antibody detected PxG2 at high levels in some ommatidia, low levels in others, and none in the remaining ommatidia. Strikingly, *spineless* knockout tissue that lacked PxBlue expression in bR1,2 showed weak PxG2 expression in all ommatidia (Fig. 4f), that is, all type II ommatidia.

In summary, CRISPR knockout of *spineless* produced ommatidia that exhibited all characterized features of type II (UV/UV) ommatidia. Therefore, Spineless expression in the two *P. xuthus* dR7-like cells controls not only PxBlue versus PxUV expression, but also coordination of the stochastic decision across the entire ommatidium, effectively specifying multiple features required for visual function. While expression of Spineless in *D. melanogaster* R7 instructs only R8 to express specific

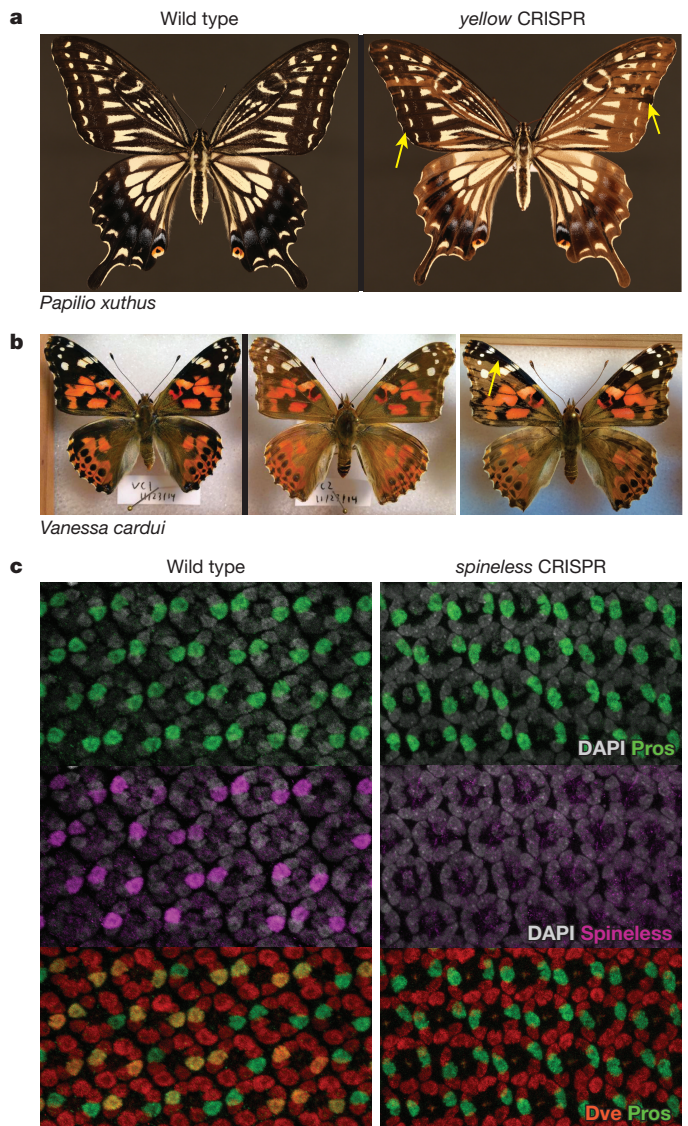


Figure 3 | Targeted CRISPR/Cas9 knockout of *yellow* and *spineless*. **a**, Wild-type *P. xuthus* are yellow or black on most dorsal surfaces (left). Knockout of *yellow* (right) causes the loss of black pigmentation. The individual shown on the right is a G0 mosaic showing large regions of lighter bi-allelic mutant tissue, with examples of darker wild-type tissue marked by arrows. **b**, The *yellow* CRISPR knockout in *V. cardui* frequently produces fully mutant G0 animals (middle, compare to wild-type at left). Mosaics are also produced (right). Results from *V. cardui* are described in Extended Data Fig. 1. **c**, Targeting *spineless* produces pupal retinas that lack Spineless protein in large domains (middle panel (right), compare to wild-type at middle panel (left)). Downstream expression of Dve matches Spineless; compare middle panel (left) Spineless (magenta) versus the bottom panel showing co-expression of Dve (red) in the same Pros-positive (green) nuclei. Knockout of *spineless* (middle panel (right)) eliminates Dve expression in Pros-positive cells (no co-expression in bottom panel (right)).

Rhodopsins⁴, *P. xuthus* dR7 cells influence the entire ommatidium, including the SVF outer photoreceptors.

The recruitment of two dR7-like cells that make independent binary stochastic choices allows for three outcomes. Interestingly, the extra dR7-like butterfly photoreceptor is recruited in precisely the same place within the developing ommatidium where a cell known as the ‘mystery cell’ has been described in *D. melanogaster*²⁷. This enigmatic cell briefly begins to take on properties of a photoreceptor cell before disappearing²⁷. It is possible that butterflies retain this extra cell during development as an additional dR7-like photoreceptor.

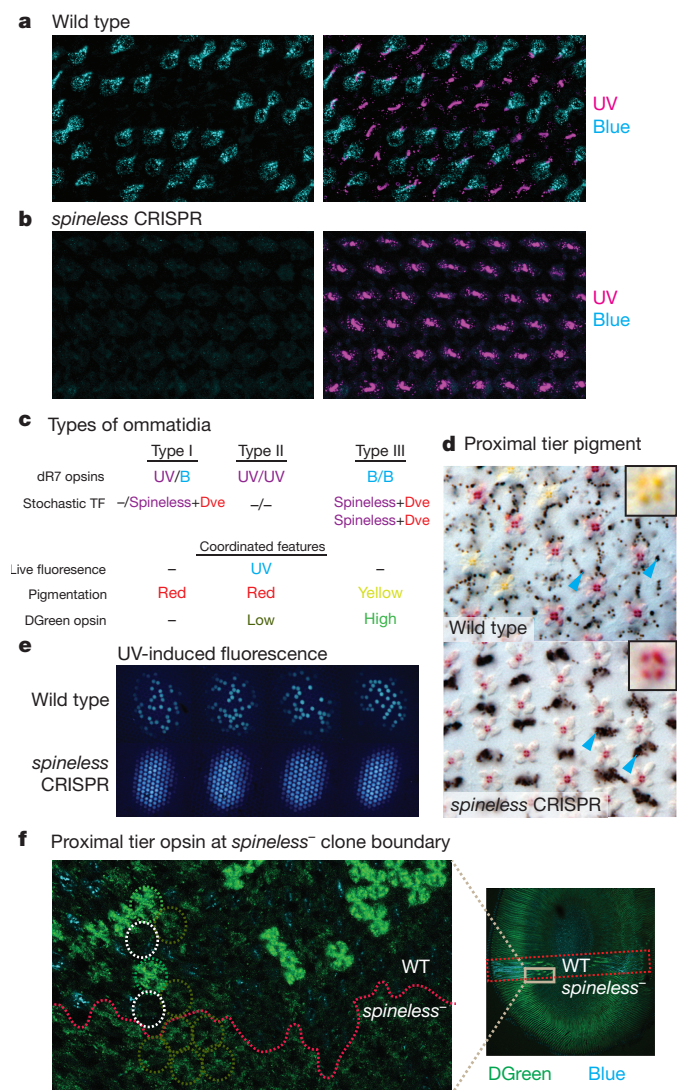


Figure 4 | Outcome of *spineless* knockout and coordination of ommatidial fate. **a**, Wild-type *P. xuthus* retinas contain a stochastic mosaic of ommatidia expressing either PxUV (magenta) or PxBlue (cyan) Rhodopsin in the dR7-like photoreceptors of the distal tier (bR1 + 2). Ommatidia are UV/B, UV/UV or B/B. **b**, In *spineless* mutant tissue, all ommatidia become UV/UV and lose blue-sensitive Rhodopsin expression. **c–f**, Other features of ommatidial type are also coordinated, as summarized in **c**. DGreen, double-peaked green Rhodopsin. **d**, Lower tier photoreceptors bR3–8 produce either a red (types I and II) or yellow pigment (type III) matching the stochastic mosaic (top). Arrowheads mark examples of dark granules of screening pigment. In *spineless* knockout retinas only red pigment is observed, indicating loss of type III (bottom). **e**, Blue fluorescence under UV light normally indicates type II fate; all ommatidia become UV fluorescing in *spineless* knockout tissues and are therefore type II. The four panels show a series along the anterior–posterior axis across one wild-type retina and one *spineless* CRISPR knockout retina. **f**, Long wavelength green-sensitive PxG2 Rhodopsin expression in the proximal tier b5–8 photoreceptors is coordinated with ommatidial type. PxG2 is highly expressed in type III, expressed at low levels in type II, and not expressed in type I. Low levels of expression in *spineless* mutant tissue indicate type II fate. The red line indicates the approximate boundary between wild-type and *spineless* mutant tissues; dashed circles indicate examples of the three ommatidial types in wild-type (top) and mutant (bottom) tissues.

On average, type I is present at 50%, type II at 25%, and type III at 25% frequencies, which corresponds to a 50% probability of Spineless expression in a given dR7-like photoreceptor (Extended Data Table 1). Previous studies found no differences between PxBlue/PxUV and PxUV/PxBlue ommatidia (Spineless ON/OFF versus OFF/ON) in

P. xuthus^{1,8}, suggesting that the two dR7 cells are equivalent before the stochastic decision to express Spineless. In *D. melanogaster*, 65% of R7 cells have Spineless^{ON}. Honeybees also have nine photoreceptors per ommatidium and three stochastically specified ommatidial types, but have a different ratio of ommatidial types^{28,29}. We predict that bees also have two dR7-like cells^{3,16,28} with a probability of 30% for Spineless expression instead of 50% in butterflies (Extended Data Table 1). The mechanisms controlling stochastic Spineless expression ratios are as yet unknown.

This work provides evidence that our extensive knowledge of patterning in the *D. melanogaster* visual system allows us to understand how vision has evolved in other insects. Expression of Sal, Pros, and Sens is deeply conserved and they are reliable markers for establishing photoreceptor homology. Stochastic patterning mechanisms involving Spineless and Dve are also shared. Adaptations for specific visual requirements can occur through modification of the developmental network that patterns the eye, such as the production of two dR7-like cells in butterflies. Importantly, the SVF photoreceptors must be able to determine whether none, one or both R7-like cells have chosen to express Spineless in order to choose which Rhodopsin(s) and pigments to express. This coordination is important for the function of each type of ommatidial unit. It will be interesting to determine whether the evolution of additional ommatidial types required coordinated changes in the brain for visual information processing.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 17 November 2015; accepted 31 May 2016.

Published online 6 July 2016.

- Arikawa, K. Spectral organization of the eye of a butterfly, *Papilio*. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **189**, 791–800 (2003).
- Briscoe, A. D. Reconstructing the ancestral butterfly eye: focus on the opsins. *J. Exp. Biol.* **211**, 1805–1813 (2008).
- Wernet, M. F., Perry, M. W. & Desplan, C. The evolutionary diversity of insect retinal mosaics: common design principles and emerging molecular logic. *Trends Genet.* **31**, 316–328 (2015).
- Wernet, M. F. *et al.* Stochastic *spineless* expression creates the retinal mosaic for colour vision. *Nature* **440**, 174–180 (2006).
- Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
- Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
- Briscoe, A. D. Six opsins from the butterfly *Papilio glaucus*: molecular phylogenetic evidence for paralogous origins of red-sensitive visual pigments in insects. *J. Mol. Evol.* **51**, 110–121 (2000).
- Koshitaka, H., Kinoshita, M., Vorobyev, M. & Arikawa, K. Tetrachromacy in a butterfly that has eight varieties of spectral receptors. *Proc. R. Soc. B* **275**, 947–954 (2008).
- Franceschini, N., Kirschfeld, K. & Minke, B. Fluorescence of photoreceptor cells observed *in vivo*. *Science* **213**, 1264–1267 (1981).
- Rister, J. & Desplan, C. The retinal mosaics of opsin expression in invertebrates and vertebrates. *Dev. Neurobiol.* **71**, 1212–1226 (2011).
- Chou, W. H. *et al.* Patterning of the R7 and R8 photoreceptor cells of *Drosophila*: evidence for induced and default cell-fate specification. *Development* **126**, 607–616 (1999).
- Johnston, R. J. Jr *et al.* Interlocked feedforward loops control cell-type-specific Rhodopsin expression in the *Drosophila* eye. *Cell* **145**, 956–968 (2011).
- Papatsenko, D., Sheng, G. & Desplan, C. A new rhodopsin in R8 photoreceptors of *Drosophila*: evidence for coordinate expression with Rh3 in R7 cells. *Development* **124**, 1665–1673 (1997).
- Kitamoto, J., Sakamoto, K., Ozaki, K., Mishina, Y. & Arikawa, K. Two visual pigments in a single photoreceptor cell: identification and histological localization of three mRNAs encoding visual pigment opsins in the retina of the butterfly *Papilio xuthus*. *J. Exp. Biol.* **201**, 1255–1261 (1998).
- Kinoshita, M. & Arikawa, K. Color and polarization vision in foraging *Papilio*. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **200**, 513–526 (2014).
- Friedrich, M., Wood, E. J. & Wu, M. Developmental evolution of the insect retina: insights from standardized numbering of the homologous photoreceptors. *J. Exp. Zool. B Mol. Dev. Evol.* **316B**, 484–499 (2011).
- Mollereau, B. *et al.* Two-step process for photoreceptor formation in *Drosophila*. *Nature* **412**, 911–913 (2001).
- Takemura, S.-Y., Kinoshita, M. & Arikawa, K. Photoreceptor projection reveals heterogeneity of lamina cartridges in the visual system of the Japanese yellow swallowtail butterfly, *Papilio xuthus*. *J. Comp. Neurol.* **483**, 341–350 (2005).

19. Xie, B., Charlton-Perkins, M., McDonald, E., Gebelein, B. & Cook, T. Senseless functions as a molecular switch for color photoreceptor differentiation in *Drosophila*. *Development* **134**, 4243–4253 (2007).
20. Cook, T., Pichaud, F., Sonnevile, R., Papatsenko, D. & Desplan, C. Distinction between color photoreceptor cell fates is controlled by Prospero in *Drosophila*. *Dev. Cell* **4**, 853–864 (2003).
21. Arikawa, K., Mizuno, S., Kinoshita, M. & Stavenga, D. G. Coexpression of two visual pigments in a photoreceptor causes an abnormally broad spectral sensitivity in the eye of the butterfly *Papilio xuthus*. *J. Neurosci.* **23**, 4527–4532 (2003).
22. Thanawala, S. U. *et al.* Regional modulation of a stochastically expressed factor determines photoreceptor subtypes in the *Drosophila* retina. *Dev. Cell* **25**, 93–105 (2013).
23. Burgess, E. A. & Duncan, I. Direct control of antennal identity by the *spineless-aristapedia* gene of *Drosophila*. *Mol. Gen. Genet.* **221**, 347–357 (1990).
24. Kitamoto, J., Ozaki, K. & Arikawa, K. Ultraviolet and violet receptors express identical mRNA encoding an ultraviolet-absorbing opsin: identification and histological localization of two mRNAs encoding short-wavelength-absorbing opsins in the retina of the butterfly *Papilio xuthus*. *J. Exp. Biol.* **203**, 2887–2894 (2000).
25. Arikawa, K. & Stavenga, D. Random array of colour filters in the eyes of butterflies. *J. Exp. Biol.* **200**, 2501–2506 (1997).
26. Arikawa, K. *et al.* An ultraviolet absorbing pigment causes a narrow-band violet receptor and a single-peaked green receptor in the eye of the butterfly *Papilio*. *Vision Res.* **39**, 1–8 (1999).
27. Tomlinson, A., Bowtell, D. D. L., Hafen, E. & Rubin, G. M. Localization of the *sevenless* protein, a putative receptor for positional information, in the eye imaginal disc of *Drosophila*. *Cell* **51**, 143–150 (1987).
28. Ready, D. F. A multifaceted approach to neural development. *Trends Neurosci.* **12**, 102–110 (1989).
29. Wakakuwa, M., Kurasawa, M., Giurfa, M. & Arikawa, K. Spectral heterogeneity of honeybee ommatidia. *Naturwissenschaften* **92**, 464–467 (2005).
30. Rister, J., Desplan, C. & Vasilaukas, D. Establishing and maintaining gene expression patterns: insights from sensory receptor patterning. *Development* **140**, 493–503 (2013).

Acknowledgements We thank members of the Desplan and Arikawa laboratories for discussion, and especially M. Wernet and J. Rister for suggestions. We thank M. Friedrich for clarifying insect eye homologies, A. Stolfi for discussing CRISPR/Cas9 protocols, C. Merlin for discussing butterfly injection technique, and J. Bothma for discussion. We thank A. Monteiro for providing anti-Sal and K. Shi at Genscript for help with antibody design. This work was supported by NIH grant EY13010 and the Center for Genomics and Systems Biology of NYU Abu Dhabi to C.D., and the JSPS Kakenhi grant numbers 26251036 and 20167232 to K.A. M.P. was supported by an NIH Ruth L. Kirschstein NRSA, a JSPS Short Term Fellowship award, and the Revson Biomedical Research Foundation Postdoctoral Fellowship.

Author Contributions M.P. and C.D. jointly conceived the project with input from K.A. M.P., C.D., and K.A. designed experiments. M.P., M.K., and L.H. performed experiments. G.S. was responsible for genomic data analysis and *de novo* assembly. C.D. and K.A. supervised the project. M.P., K.A., and C.D. wrote the paper. All authors discussed the results and manuscript.

Author Information Data was deposited at SRA under accession code SRP071804. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.D. (cd38@nyu.edu).

METHODS

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Animals. *P. xuthus* were reared in the laboratory in Kanagawa, Japan using a line derived from wild-caught females. The larvae were fed on fresh citrus leaves under a light regime of 14 h light and 10 h dark at 28 °C; these conditions produce non-diapausing spring form adults. Adults eclosed approximately 10 days post-pupation.

V. cardui and artificial diet for larval culture were obtained from Carolina Biological Supply. Larvae were reared individually in disposable plastic 1 oz containers. Adults were maintained on sugar water and fruit slices in desktop population cages under artificial light. Egg collections were performed using stems of cut sunflowers.

Antibody generation. Proteins were produced by GenScript (Piscataway, NJ) and purified to > 80% purity. Codon optimized gene synthesis was used to generate the sequences of interest for expression in bacteria. GenScript performed protein injections into host animals, collected serum for testing, and performed affinity purification for each antibody (except PxDve). Antibodies were tested at a range of concentrations and fixation conditions in freshly fixed eye and embryo tissue. Embryo staining patterns provided an indication of whether staining might be specific; early transcription factor expression patterns are often highly conserved and stains showed roughly *D. melanogaster*-like patterns. Staining in the eye in subsets of the appropriate cell types was used as another indication; we show that these regulatory networks are highly conserved and the staining patterns match morphological features as well as each other (for example, Dve correlates with Spineless in R7-like photoreceptors). Further evidence of antibody specificity was provided by targeted knockout experiments showing loss of expression patterns of the target (Spineless) or by changes in expression patterns of downstream targets (Dve, Rhodopsins). Rhodopsin expression patterns had been previously determined via *in situ* hybridization and rhodopsin antibodies were confirmed to be mutually exclusive and in the correct patterns.

Sequences used for protein production were PxDve: SHSPNSGLMASN DNYAMQDRNSMKSPMQMSGSPGRYPMSIMSEDNLSNAGSDLEDDGGDLN PDDRPESPDAPLSLITTKNNNNDEEMSKQSPKPPDIKVHDIQELRLDTRSD QTNSPQRSPDKNSDSSHHNNNNNVKEENGITDEQDVASDDDIVQERHYRPS TPHLDRLPFPMVFNHPMFHGMIMSYMTGFGVGPVPGEGASGLNLA LAGASDERKRNRFTIDPVSEVPVLEQWFSMNTHPSHNLILKYTEELNRMYP RQKFPRLSKNVQWFKNRRACKRLKMSLYEPSSSHYSHPGHHAIA; PxPros: PMHLRKAALMFVWVRYPSSAVLKMYPFDIKFNKNNTAQLVKWFSNREFYY IQMEKYARQAISEGLKAADDLHVAGDSELYRVLNLHYNRNNHIEVPPNFY VVEQTLREFFRAIQGGKDTQESWKSIIYKVISRLDDPVPEYFKSPNLEQLE; PxSs: GAYTGPYGEYPPAPALHYAPPLDDRFLAADNLHFQYKPLPYHYTPYA PNGFLEPAPPPGYEVAPTYHRPPSREYTYVDSTGRYMSPTVGQERRSPSVVP GPSPGGSSGSTEQDRLQQTQSPSEIPRQTVLMWGAGGTAQEVPEYSPQVW RYHQSHYHTAEATQ; VcSens: GVDHARHTPPRDDDEELPLNLS MKNRQIWPSPASVQWERESPMSRWDAHDDPSLELVKRCRNSNDE KRPSSTEPLRCPAPAYHSYPPPPPPADINISLLKTENKNEKSFQCKQCGKCF KRSSTLSTHLLIHSSTRPYPCQYCGKRFHQKSDMKKHTYIHTGEKPHKCVV CSKAFSSQSNLITHMRKHTGYKPFSCGLCDKAFQRKVDLRRHRESQHSVD PPSIALSQNRIFYGDDQTSAPISN; PxBlue: INHPRYAELQKRLPC; PxUV: CISHPKYQELQKRM; PxG2: CAIANLEPGMGASEA.

Immunofluorescence. Antibody staining was performed as described previously for *D. melanogaster* pupal retinas³¹. The following antibodies concentrations were used: guinea pig anti-Sal^{32,33} at 1:10,000 (gift of Antonia Monteiro), and the antibodies generated in this work were: rat anti-PxPros 1:100, rabbit anti-VcSens 1:300, rabbit anti-PxSs 1:100, guinea pig anti-PxDve 1:400, rabbit anti-PxBlue 1:400, guinea pig anti-PxUV 1:100, and rabbit anti-PxG2 1:100. Pupal retina stains were performed at 2–6 days post-pupation, with ~72 h post-pupation providing the most optimal staging. Images were captured using standard confocal microscopy on a Leica SP5.

CRISPR/Cas9 somatic mutagenesis. We used T7 transcription-based *in vitro* reactions to generate single guide RNAs (sgRNAs) from target-specific gBlock DNA fragments ordered from Integrated DNA Technologies. We used a modified 'F + E' sgRNA scaffold to prevent premature transcriptional termination³⁴. These sgRNAs were injected with recombinant Cas9-NLS protein (PNA Bio).

CRISPR sgRNAs. Target sequences were identified by eye and tested for uniqueness by BLASTing against the respective transcriptome and genome sequences. Regions that did not contain polymorphism at the target site were preferred. Target sequences used are as follows, with 'Px' for *P. xuthus* and 'Vc' for *V. cardui*: Px_y1_gRNA: GACAGATGCGACAGACTC; Px_y2_gRNA: GGTACACTGTGGTGAGTC;

Px_ss2_gRNA: CAACATTTTCAGGAAGAGG; Px_ss3_gRNA: CACATACC AGGAATCCAG; Vc_y1_gRNA: CTCTGACGAGCTCGGCTA; Vc_y2_gRNA: GCCTTATCGCTTACTCCT; Vc_ss1_gRNA: TAAGAGACGACGGAGG AA; Vc_ss2_gRNA: TCTCGGTGCCTCTTGCTC. These targets were incorporated into the following scaffold sequence in place of the Ns: GTAAA ACAGCGGCCAGTGAATTGTAATACGACTCACTATA gNNNNN NNNNNNNNNNNNNGTTTAAAGAGCTATGCTGGAAACAGCATAGCAAGT TTAAATAAGGCTAGTCCGTATCAACTGAAAAAGTGGCACCAGAGTCGG TGCTTTTTTTTGTATAGCTGTTTCTCTG. This sequence is composed of: M13_spacer_T7 promoter_target_scaffold_M13R. The region of target-specific sequence is underlined.

The first base pair of the target sequence was in each case changed to a G for efficient *in vitro* T7 transcription; the sequence at this 5' position is not required for targeting. The scaffold sequence used was from ref. 34. M13F and M13R sequences flank the fragment and are useful if additional amplification is necessary. M13F is followed by a short spacer sequence and the T7 promoter.

Preparation of injection constructs. Templates were ordered as 'gBlocks' (synthesized DNA fragments) from Integrated DNA Technologies. These were PCR amplified using M13F and an internal reverse primer AAAAAAGCACCAGCTCGGTGCC. For a given sgRNA, products from four 50 µl PCR reactions were pooled and purified on single Qiagen spin columns. Resulting template was used for *in vitro* transcription using the AmpliScribe T7-flash Transcription Kit from Epicentre/Illumina to produce sgRNAs and followed by ammonium acetate precipitation for purification. Products were resuspended in Qiagen buffer EB, quantified by spectrophotometry using a NanoDrop, and stored at –80 °C until injection. They were then mixed with Cas9-NLS protein from PNA Bio (catalogue number CP01) at a final concentration of 250–500 ng µl^{–1} of sgRNA and of Cas9 protein.

Injections. Embryo injections were performed using a Narshige micromanipulator and a 50 ml syringe with pressure applied by hand. Needles were pulled using a Sutter Instrument micropipette puller and needle tips were broken by hand using forceps against a glass slide. Needles were front-loaded under the microscope with 0.8 µl volumes of injection construct.

Embryos were collected for 0–4 h post laying and either injected immediately for high percentage, biallelic somatic knockout or aged until 5–8 h after eggs were laid for generation of mosaic animals. Eggs were surface sterilized using ~5 min submersion in 8% benzalkonium chloride solution, which also helps remove the 'glue' used to secure eggs upon laying. Eggs were then briefly rinsed and transferred to a glass microscope slide using a damp paint brush, and blunt forceps were used to position embryos onto thin strips of double-sided tape. After injection, embryos were placed into humid Petri dishes until hatching (~4 days). Hatchlings were collected via paintbrush and transferred to leaves or artificial diet.

Yellow mutation is apparent in early larval head capsules. Wild-type head capsules in both *P. xuthus* and *V. cardui* are black, with mutant tissue appearing light brown or tan. In *V. cardui*, larval spines that are normally black in instars 3–4 become lighter in coloration in Y– tissue. Spineless mutation has no apparent effect until pupation when mutant antennae provide indication of successful *spineless* knockout. Variability in the severity of antennal defects produced provide an early estimate of knockout efficiency. One *yellow* mutation was isolated and maintained as a viable line for several generations.

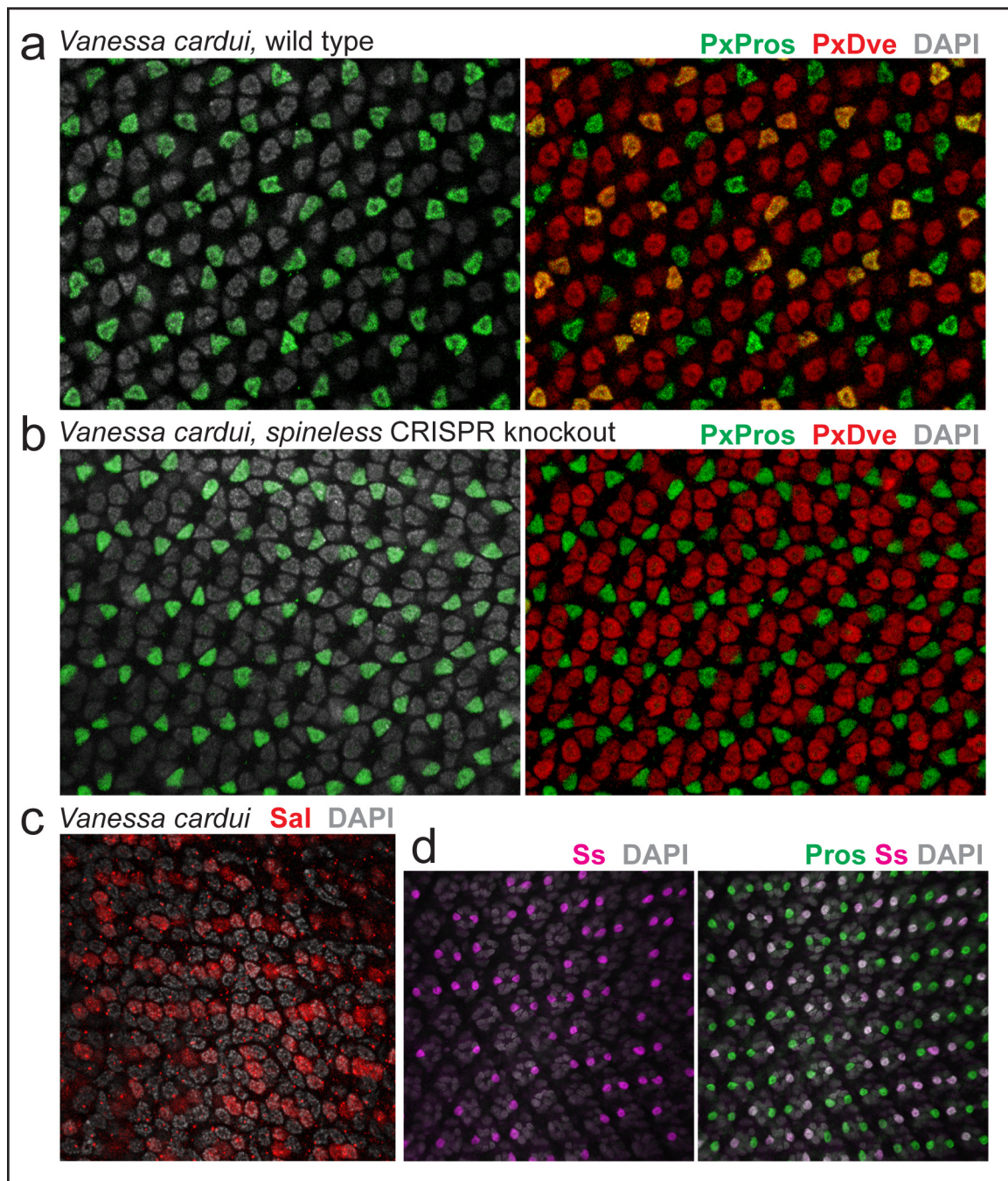
Mutation sequencing. DNA was isolated from tissue showing appropriate phenotypes using the Qiagen DNeasy kit. PCR products were cloned using the pGEM-T Easy Vector system and sequenced with M13R primers. Examples of the resulting sequences are presented in Extended Data Fig. 2e. The primers used for PCR amplifications were 565VcSs1MutSeqF1: ATACGAGCCAGCTTGGTACC; 566VcSs1MutSeqR1: GTCTTTTGATGGTGGCTTG; 567VcSs2MutSeqF1: TCAAGCCACCATCAAAAGAC; 568VcSs2MutSeqR1: GACCACCTG AAAATAGCTC; 569VcYMutSeqF1: CTTTAGACGCTCCCTATGAAC; 570VcYMutSeqR1: ATGTTGAAGTCTCCGACAAG; 595VcSs1F2: AGTTACAATTGTGCGGCTG; 596VcSs2R2: TGGTCGTAAATAGCTGACG.

Results from *V. cardui*. Results from *V. cardui* are summarized in Extended Data Fig. 1. *V. cardui* also have two photoreceptors per ommatidium labelled with Pros, three per ommatidium labelled with Sal, and a stochastic mosaic of Spineless and Dve expression. Knockout of Spineless via CRISPR/Cas9 produces fields of uniform Dve-negative R7s, as in *P. xuthus*. There are no known coordinated features of ommatidial type in *V. cardui*; they express a single green Rhodopsin in all outer SVF photoreceptors and have uniform pigmentation. This lack of coordinated features such as red/yellow screening pigments or differential Rhodopsin expression in the SVF, proximal tier prevented an assessment of ommatidial coordination in *V. cardui*. However, Spineless expression clearly causes loss of downstream Dve expression, as in *P. xuthus*, and therefore does account for known features of

ommatidial type. Spineless expression is presumably coordinated with the UV/UV type (of the three UV/UV, UV/B, and B/B types), as in *P. xuthus*, though this was not tested directly in *V. cardui*.

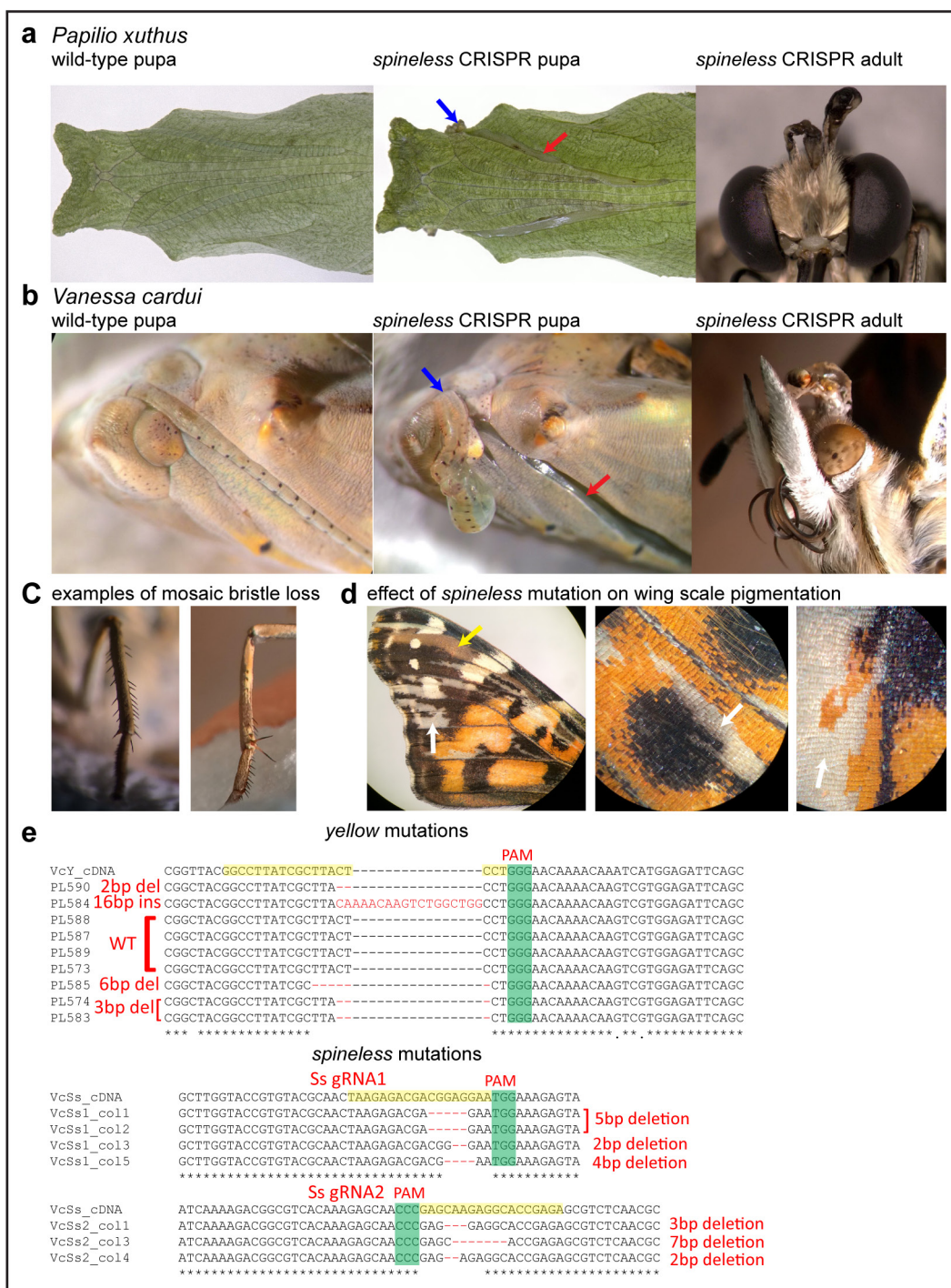
Adult morphology and staining. Yellow and red pigment was viewed using light microscopy of plastic sections, as described in ref. 35. UV fluorescence was assessed by mounting live animals on a goniometer for viewing under UV illumination. Rhodopsin antibodies were initially tested on hand-dissected fragments of adult retina tissue. This approach did not provide clear images of spatial patterns and so a thick-sectioning approach was used. Corneal lenses were hand-removed from dissected retinas after an initial fixation as described in³⁶. Retinas were then embedded in gelatin and cut into 100 μ M sections, as in ref. 36, with the overnight fixation concentration adjusted from 8% to 6%. Sectioned tissue was then stained following standard protocols³¹.

31. Hsiao, H. Y. *et al.* Dissection and immunohistochemistry of larval, pupal and adult *Drosophila* retinas. *J. Vis. Exp.* <http://dx.doi.org/10.3791/4347> (2012).
32. Stoehr, A. M., Walker, J. F. & Monteiro, A. Spalt expression and the development of melanic color patterns in pierid butterflies. *Evodevo* **4**, 6 (2013).
33. de Celis, J. F., Barrio, R. & Kafatos, F. C. Regulation of the Spalt/Spalt-related gene complex and its function during sensory organ development in the *Drosophila* thorax. *Development* **126**, 2653–2662 (1999).
34. Chen, B. *et al.* Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491 (2013).
35. Arikawa, K., Scholten, D. G. W., Kinoshita, M. & Stavenga, D. G. Tuning of photoreceptor spectral sensitivities by red and yellow pigments in the butterfly *Papilio xuthus*. *Zoolog. Sci.* **16**, 17–24 (1999).
36. Kinoshita, M., Shimohigashi, M., Tominaga, Y., Arikawa, K. & Homberg, U. Topographically distinct visual and olfactory inputs to the mushroom body in the Swallowtail butterfly, *Papilio xuthus*. *J. Comp. Neurol.* **523**, 162–182 (2015).



Extended Data Figure 1 | Additional expression data and CRISPR/Cas9 *spineless* knock out results for *V. cardui*. **a**, Antibodies to *P. xuthus* Pros and Dve (PxPros and PxDve) cross-react in *V. cardui*. As in *P. xuthus*, Pros labels two R7-like photoreceptors per ommatidium (green) while Dve labels bR3–8, the SVF photoreceptors equivalent to the outer photoreceptors of *D. melanogaster* (red), along with a stochastic subset of

Pros-expressing R7-like PRs (yellow co-expression). **b**, Co-expression of Dve in Pros-positive R7-like photoreceptors is lost in *spineless* CRISPR/Cas9 knockout tissue; compare to yellow co-expression in wild-type in **a**. **c**, Sal antibodies label three photoreceptors per ommatidium in *V. cardui* (red). **d**, As in *P. xuthus*, Spineless antibodies (magenta) label a stochastic subset of R7-like Pros-expressing photoreceptors (green).



Extended Data Figure 2 | Pleiotropic effects of the *spineless* mutation and sequencing of CRISPR/Cas9 generated mutations. a, b, The *spineless* mutation affects antennal development in *P. xuthus* (a) and *V. cardui* (b), as shown previously for *D. melanogaster*. This effect is first visible at pupal stages, where missing/shortened antennae are absent in their normal channels (red arrow) and shortened structures protrude in their place (blue arrow). Strongly affected individuals have little to no remaining antennae in adult stages (at right). c, The *spineless* mutation affects bristle development in *V. cardui*, as shown previously for *D. melanogaster*, where some bristles are missing or reduced. In the left image, the regular comb of bristles on the tibia is interrupted in mosaic *spineless* mutant adults, or almost completely missing (right image). d, Mutation of *spineless*

produced an unexpected effect on wing colour pattern. The wing of an animal injected with guide RNAs targeting both *spineless* and *yellow* during the late blastoderm stage (7–9 h after egg lay) is shown on the left. A brown-coloured *yellow* mutant patch of tissue is visible (yellow arrowhead), as shown in the *yellow* mutation in Fig. 3, but lighter patches of wing scales lacking both melanin (black) and ommachromes (oranges) are also visible, example marked with a white arrow. Similar clones were observed when only *spineless* sgRNAs were injected (middle and right). This effect was observed for two independent sgRNAs targeting *Spineless*. e, Cloning and sequencing of target regions from mutant tissues reveals a mixture of CRISPR/Cas9 generated mutations. Unmodified cDNA sequences are shown on the top line for comparison.

Extended Data Table 1 | Specific probabilities of Spineless expression yield specific ratios of ommatidial type

Given a pair of R7 cells with independent probabilities to express Ss, there are four possible combinations (cell1; cell2):

Ss+; Ss+ $p1 \cdot p2$

Ss+; Ss- $p1 \cdot (1-p2)$

Ss-; Ss+ $(1-p1) \cdot p2$

SS-; Ss- $(1-p1) \cdot (1-p2)$

If we assume that $p1=p2$ and make $p=p1=p2$ then:

$p \cdot p$

$p \cdot (1-p)$

$p \cdot (1-p)$

$(1-p) \cdot (1-p)$

Probability	Type II	Type Ia	Type Ib	Type III	
0.1	0.01	0.09	0.09	0.81	
0.2	0.04	0.16	0.16	0.64	
0.3	0.09	0.21	0.21	0.49	
0.4	0.16	0.24	0.24	0.36	
0.5	0.25	0.25	0.25	0.25	
0.6	0.36	0.24	0.24	0.16	
0.7	0.49	0.21	0.21	0.09	
0.8	0.64	0.16	0.16	0.04	
0.9	0.81	0.09	0.09	0.01	
With on/off and off/on combined:					
Probability	Type II	Type I	Type III		
0.1	0.01	0.18	0.81		
0.2	0.04	0.32	0.64		
0.3	0.09	0.42	0.49		
0.4	0.16	0.48	0.36		
0.5	0.25	0.5	0.25	observed for <i>Papilio</i>	
0.6	0.36	0.48	0.16		
0.7	0.49	0.42	0.09	observed for <i>Apis</i>	
0.8	0.64	0.32	0.04		
0.9	0.81	0.18	0.01		

If Spineless has a specific, single probability of being expressed and if the two R7-like cells make this decision independently then a specific ratio of the three ommatidial types will be produced. These ratios are shown for example probabilities. If ON/OFF and OFF/ON (type Ia and type Ib) are equivalent, these types can be combined, shown below. The ratio observed for *P. xuthus* supports the idea that the two R7-like cells are independent and, on average, have a 50% probability of Spineless expression. Interestingly a different ratio that matches this model has previously been observed for the honeybee *Apis mellifera* (see Discussion section).

Dissociated functional significance of decision-related activity in the primate dorsal stream

Leor N. Katz^{1*}, Jacob L. Yates^{1*}, Jonathan W. Pillow² & Alexander C. Huk¹

During decision making, neurons in multiple brain regions exhibit responses that are correlated with decisions^{1–6}. However, it remains uncertain whether or not various forms of decision-related activity are causally related to decision making^{7–9}. Here we address this question by recording and reversibly inactivating the lateral intraparietal (LIP) and middle temporal (MT) areas of rhesus macaques performing a motion direction discrimination task. Neurons in area LIP exhibited firing rate patterns that directly resembled the evidence accumulation process posited to govern decision making^{2,10}, with strong correlations between their response fluctuations and the animal's choices. Neurons in area MT, in contrast, exhibited weak correlations between their response fluctuations and choices, and had firing rate patterns consistent with their sensory role in motion encoding¹. The behavioural impact of pharmacological inactivation of each area was inversely related to their degree of decision-related activity: while inactivation of neurons in MT profoundly impaired psychophysical performance, inactivation in LIP had no measurable impact on decision-making performance, despite having silenced the very clusters that exhibited strong decision-related activity. Although LIP inactivation did not impair psychophysical behaviour, it did influence spatial selection and oculomotor metrics in a free-choice control task. The absence of an effect on perceptual decision making was stable over trials and sessions and was robust to changes in stimulus type and task geometry, arguing against several forms of compensation. Thus, decision-related signals in LIP do not appear to be critical for computing perceptual decisions, and may instead reflect secondary processes. Our findings highlight a dissociation between decision correlation and causation, showing that strong neuron–decision correlations do not necessarily offer direct access to the neural computations underlying decisions.

We investigated the functional significance of decision-related activity by recording and inactivating neural activity in two well-studied cortical areas, MT and LIP, while rhesus monkeys performed a challenging motion discrimination task. On each trial, the monkey maintained stable visual fixation while discriminating the net direction of a visual motion stimulus, and then made a saccade to one of two choice targets to communicate their choice (Fig. 1a, b). For electrophysiological recordings in MT, we placed the motion stimulus in the receptive field of the neurons and aligned it with the preferred direction of one or more MT neurons on the multi-electrode array. For LIP, we placed one of the two targets in the response field of the neurons (as opposed to the visual motion stimulus), and the other target on the contralateral side of the visual field, consistent with previous studies of decision-related responses in LIP¹¹.

We recorded 157 MT neurons and 200 LIP neurons with either single electrodes or multi-electrode linear arrays. MT neurons that were well-targeted by the stimulus ($n = 94$) had average firing rates that depended on the motion strength and direction (Fig. 1c). As expected

in this area, responses increased sharply after motion onset and maintained a robust firing rate throughout motion viewing¹². The average responses of well-targeted LIP neurons ($n = 113$) were also consistent with classical observations^{2,11}, exhibiting ramp-like increases or decreases in firing rate, whose slopes were proportional to motion

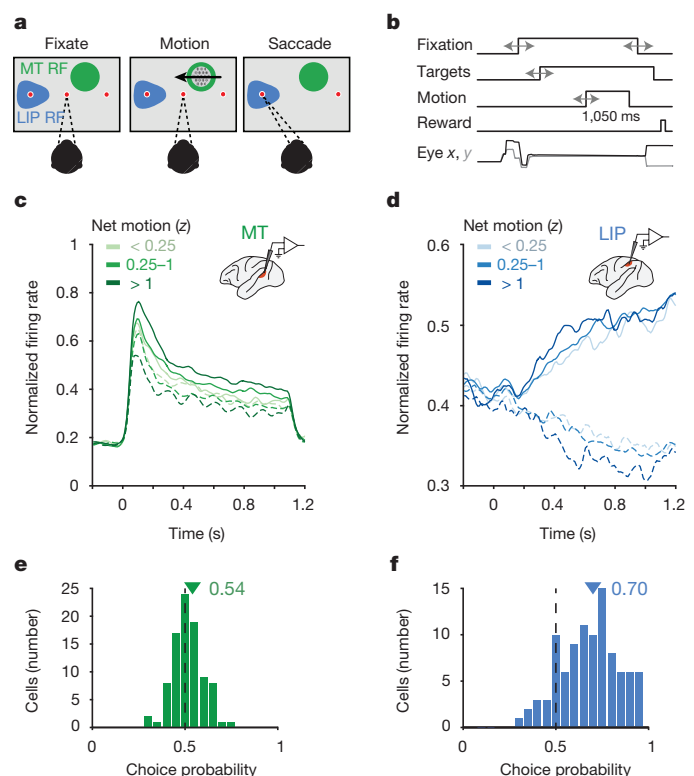


Figure 1 | Task and neural responses during direction discrimination.

a, Monkeys were trained to discriminate the direction of visual motion and communicate their decision with a saccadic eye movement to one of two choice targets. For MT recordings, motion was placed in the MT receptive field (RF) (green patch). For LIP recordings, one of the saccade targets was placed in the LIP receptive field (blue patch). **b**, Sequence of task events. Grey arrows indicate temporal jitter. **c**, Average response of 94 MT neurons as a function of motion strength (grouped by z -scored net motion, see Methods) and direction (preferred versus non-preferred direction, solid and dashed lines, respectively), aligned to motion onset. **d**, Average response of 113 LIP neurons as a function of motion strength and direction (in versus out of cell's receptive field, solid and dashed lines, respectively), aligned to motion onset. **e**, Choice probability for 90 MT neurons computed during the motion epoch. Triangle indicates mean, 0.54. **f**, Choice probability for 96 LIP neurons computed during the motion epoch. Triangle indicates mean, 0.70. Only neurons with >20 repeats of identical stimuli were included in the choice probability analysis.

¹Center for Perceptual Systems, Departments of Neuroscience & Psychology, The University of Texas at Austin, Austin, Texas 78712, USA. ²Princeton Neuroscience Institute & Department of Psychology, Princeton University, Princeton, New Jersey 08540, USA.

*These authors contributed equally to this work.

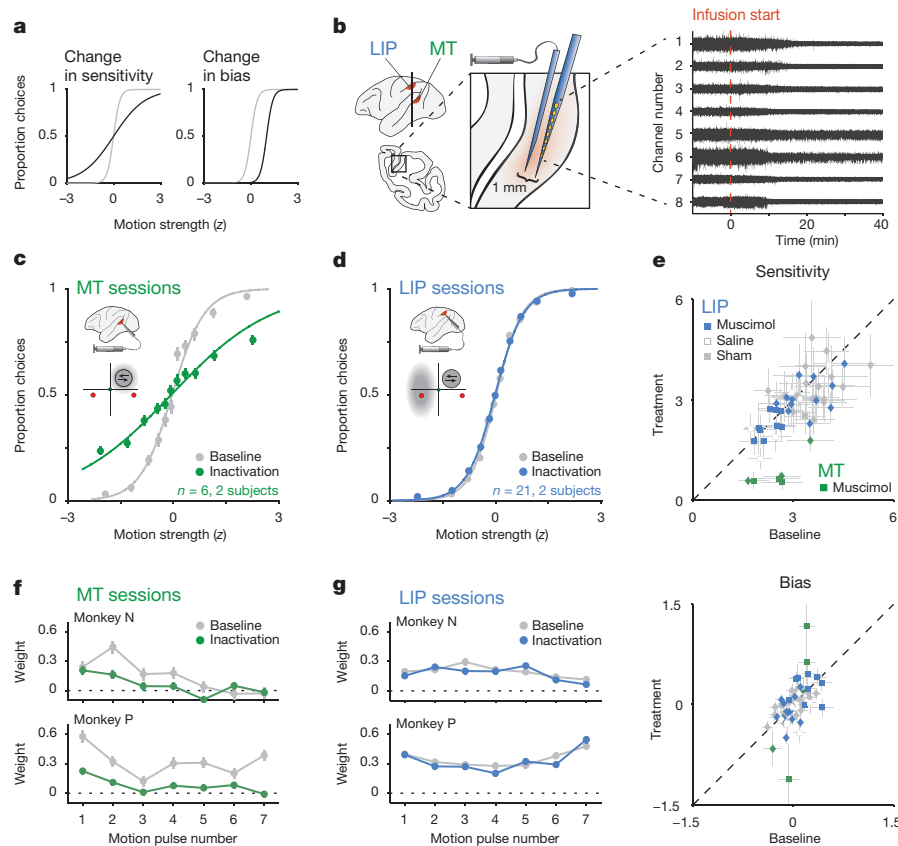


Figure 2 | Psychophysical performance before and after neural inactivations in areas MT and LIP. **a**, Hypothesized consequences of inactivation on the psychometric function. Left, decreased psychophysical sensitivity would correspond to a decrease in slope. Right, changes in psychophysical bias would correspond to a shifted midpoint. Positive values in the *x* axis (*z*-scored motion strength) refer to motion towards the target contralateral to the LIP under study. Correspondingly, the *y* axis refers to the proportion of contralateral target choices. This convention is maintained throughout. **b**, Schematic of the inactivation protocol. Left, a multi-electrode array was lowered alongside the cannula to identify the targeted cortical location, to verify neural selectivity before infusion, and to confirm neural silencing after. Right, continuous voltage traces from an example inactivation session in which neural silencing is evident ~10 min after infusion start. **c**, **d**, Psychophysical data for averaged pairs of baseline and muscimol treatment sessions in MT (**c**), and LIP (**d**). Insets illustrate

the brain region inactivated (top) and the corresponding experimental geometry (bottom), along with the estimated inactivated field (grey cloud). Error bars on points show ± 1 s.e.m. over all trials.

e, The distribution of psychometric function parameters, slope (top) and shift (bottom), reflecting sensitivity and bias, respectively, for baseline (*x* axis) and treatment (*y* axis) session pairs for MT inactivations (green symbols), LIP inactivations (blue symbols), as well as LIP saline (open grey symbols) and sham/control experiments (filled grey symbols), for monkey N (diamonds) and monkey P (squares). Error bars show 95% confidence intervals for individual sessions. **f**, **g**, Psychophysical weighting, estimated via reverse correlation. The *y* axis indicates how much the subject weighed each of the motion stimulus pulses over all baseline and inactivation session pairs in MT (**f**) and in LIP (**g**), for monkey N (top) and monkey P (bottom). Error bars are s.e.m. over all trials.

strength, the primary physiological characteristic implicating LIP in reflecting the accumulation of evidence over time (Fig. 1d).

We further quantified the decision-related activity of MT and LIP using choice probability¹, a measure of correlation between neural activity and choices, independent of stimulus-driven responses. MT neurons were weakly but reliably correlated with the animal's choice on a trial-by-trial basis (mean choice probability = 0.54, $P = 1 \times 10^{-5}$; Fig. 1e). LIP neurons were more strongly correlated with choices (mean choice probability = 0.70, $P = 1 \times 10^{-21}$; Fig. 1f). Thus, the stimulus-dependent responses and choice probability in MT were consistent with its well-established role in representing the motion stimulus, and the response patterns in LIP resembled the time course of an evolving decision process. Together, these properties have given rise to a model where LIP neurons either integrate, or reflect the integration of, motion evidence from area MT in favour of a decision^{11,13}.

Having confirmed the neurophysiological properties of areas MT and LIP and their differential degrees of correlations with decisions, we tested their respective causal contributions by performing reversible inactivations in each area and evaluating the impacts on psychophysical performance (hypothesized outcomes shown in Fig. 2a). We infused muscimol (a GABA_A agonist which hyperpolarizes cell

bodies but not fibres of passage¹⁴) into either MT or LIP, 1 mm away from a multi-electrode array (Fig. 2b). The injection cannula was targeted to locations that had yielded the largest number of canonical MT or LIP units during recording sessions (Extended Data Fig. 1). The multi-electrode array was used to confirm both pre-infusion physiological properties and post-infusion neural silencing, performed on every inactivation session. Silencing was typically observed across all recording channels of the array (Fig. 2b) and estimated to span a spherical volume of ~2.5 mm radius (see Methods).

Inactivations in area MT exerted large effects on psychophysical performance. The motion stimulus was placed within a region of visual space retinotopically matched to the inactivated population of MT neurons (Fig. 2c). MT inactivations ($n = 6$; 3 in monkey N; 3 in monkey P) had a large and consistent impact on direction discrimination sensitivity (68.5% reduction from baseline, $t_{(5)} = -9.7$, $P = 0.002$, paired *t*-test). When the motion stimulus was moved outside the inactivated region within the same session ($n = 3$), psychophysical performance was restored, confirming that the effects were not due to general changes in arousal or vigilance (Extended Data Fig. 2). These severe and specific impairments in direction discrimination performance were consistent with prior causal perturbations^{15,16}.

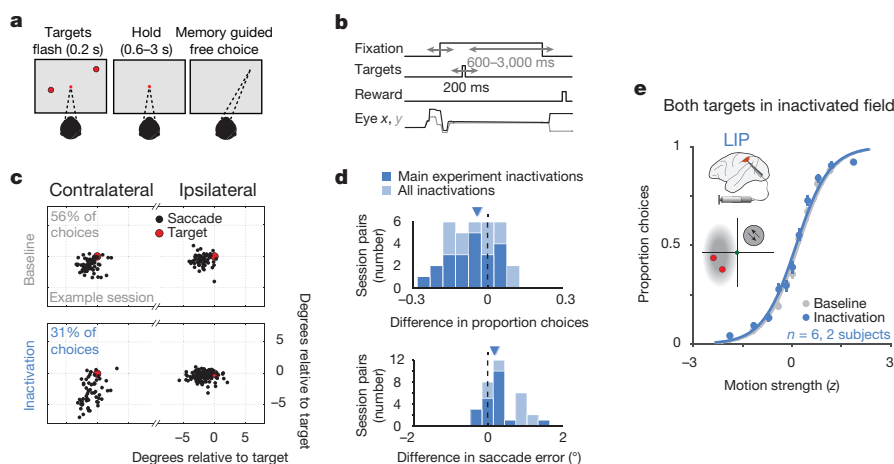


Figure 3 | Performance in control tasks following LIP inactivation.

a, Structure of the free-choice task. Following a 200-ms long presentation of two targets at random locations in space, monkeys were required to hold fixation for another 600–3,000 ms, and then to move their eyes to the remembered location of either target. **b**, Event timing in the free-choice task. Events in the task were presented in sequence and were jittered in time (grey arrows). **c**, The effect of LIP inactivation on choice bias and saccade accuracy in the free-choice task (example session). Saccade landing points (black dots) have been aligned to target position (red dot), for contralateral (left) and ipsilateral target choices (right), during baseline (top) and inactivation (bottom). Both saccadic accuracy and percent contralateral choices (indicated, top left) are reduced after LIP inactivation, to the contralateral hemifield. **d**, The effect of LIP inactivation on choice bias and saccade accuracy in the free-choice task,

over all sessions. Histograms show baseline/inactivation differences in proportion contralateral choices (top) and saccade error (bottom), where positive numbers indicate an increase in metric following inactivation. Dark bars indicate sessions that took place on the same days as the main direction discrimination experiment (main experiment inactivations, $n = 21$; 12 in monkey N; 9 in monkey P); dark triangle indicates the median difference. Light bars include an additional 13 sessions that took place during other inactivation experiments under similar conditions (all inactivations, $n = 34$; 14 in monkey N; 20 in monkey P); light triangle indicates median difference (visually occluded by dark triangle). **e**, Psychophysical data for pairs of baseline and muscimol treatment in LIP when both choice targets were placed within the inactivated field. Inset presents stimulus geometry and estimated inactivated field. Error bars are s.e.m. over all trials.

In contrast, inactivations in area LIP ($n = 21$; 12 in monkey N; 9 in monkey P) did not exert compelling or substantial effects on psychophysical performance (Fig. 2d). In these experiments, we placed one choice target in the inactivated region of visual space, consistent with previous electrophysiological investigations that placed a choice target (and not the visual motion stimulus) in the response fields of LIP neurons to elicit the area's canonical decision-related responses. Although we performed large inactivations in locations where LIP electrophysiology had mirrored the accumulation of evidence and demonstrated strong decision-related activity, we did not detect significant changes in either the animal's sensitivity or bias, as indicated by statistically indistinguishable differences in the slope (3.7% reduction from baseline, $t_{(20)} = -1.4$, $P = 0.16$, paired t -test) or midpoint (-0.4% shift, $t_{(20)} = -0.08$, $P = 0.93$, paired t -test) of the psychometric functions. Saline and sham control experiments showed similar patterns to the main baseline versus muscimol treatment comparison (Extended Data Table 1). Thus, although the effect of MT inactivation on sensitivity was substantial, an effect of LIP inactivation was not clearly identifiable using our techniques and task (Fig. 2e).

We also assessed whether inactivation affected the timing or strategy of evidence integration^{8,17,18}. For example, if LIP supported the temporal integration of motion evidence, inactivation could alter the strategy to reflect 'leakier' integration that might still support the same overall performance. Contrary to this possibility, inactivation in LIP did not lead to greater reliance on either early or late information (Fig. 2f, g), as estimated via reverse correlation. Inactivations in area MT, in contrast, reduced the psychophysical weighting of motion roughly evenly over time.

Although inactivation in LIP had no measurable effect on direction discrimination, it did exert effects on a 'free-choice' control task, which was performed on every inactivation session (Fig. 3a, b). LIP inactivation biased choices away from the contralateral hemifield (8.88% reduction from baseline on average, $t_{(33)} = 3.4$, $P = 0.001$, paired t -test), (Fig. 3c, d), consistent with previous reports in monkeys^{19–21}, rodents⁸, and parietal lesions in humans²². Thus, our electrophysiological

confirmation of LIP inactivation was complemented by a behavioural consequence in this free-choice control task. In addition to causing a spatial choice bias, LIP inactivation led to an increase in endpoint error of saccades made to the hemifield contralateral to inactivation (0.36° on average, $t_{(33)} = 4.4$, $P = 7 \times 10^{-5}$, Fig. 3c, d). No systematic change was detected in other oculomotor metrics during the free-choice task (reaction time, peak velocity, or duration), and no effects on any oculomotor metrics were detected during the direction discrimination task. Despite observing a muscimol-induced effect in the free-choice task, effect magnitude in the free-choice task was not predictive of effect magnitude in the direction discrimination task (Extended Data Fig. 3a, b), nor was there a dose–response relationship between muscimol mass and behavioural performance (Extended Data Fig. 3c–e), suggesting that our large muscimol administrations were probably operating within a 'ceiling' regime.

Because muscimol inactivations require comparisons across relatively long time scales, it remains logically possible that LIP normally plays a critical role in decision making, given that other areas are processing information in parallel and are able to quickly compensate when it is artificially inactivated. Although other techniques with faster time scales will allow for more direct tests of this possibility, we did not observe changes indicative of compensation either within a session or over sessions (Extended Data Figs 4 and 5, respectively). We also tested for compensation involving the non-inactivated hemisphere²³. We performed 6 additional inactivation experiments with both choice targets placed in a single hemifield (Fig. 3e, inset), in order to maximize reliance on a single hemisphere's LIP and hence minimize involvement of the other hemisphere. Inactivation of the LIP corresponding to the two targets did not produce clear changes in behavioural performance (Fig. 3e), indicating that inter-hemispheric compensation was unlikely in our main experiments. Previous LIP inactivation studies also find no evidence in support of compensation that manifests behaviourally (see the section on spatial and temporal extent of inactivation in the Methods). We also found no disruption of decision-making performance using the moving-dot stimulus used in

previous studies of MT and LIP function during decision making^{2,15} (Extended Data Fig. 6c).

Our results reveal a dissociation between decision-related activity in LIP and the causal role of such activity in decision making. Instead, decision-related signals in LIP may be a result of feedback²⁴, or an emergent phenomenon driven by extensive training²⁵. Although one prior study observed effects of LIP microstimulation in a reaction time direction discrimination task²⁶, such electrical perturbations can produce orthodromic (and antidromic) activation of connected areas, and their observed effects are reconcilable with multiple alternatives to evidence accumulation⁶. It remains possible that LIP contributes to decision making in conjunction with associated brain regions, whose parallel and/or redundant processing simply renders LIP unnecessary in the particular tasks used to study its decision-related activity. Indeed, a growing body of work has observed decision-related activity in other brain areas^{3–6,9}, consistent with the prospect of LIP playing a minor and/or nonessential role in decision making. Our results mirror findings in rodent posterior parietal cortex, where inactivations did not affect decision making despite electrophysiological correlates of evidence accumulation⁸. Finally, a richer appreciation of LIP's contributions to decision making might be gleaned from placing the motion stimulus itself (as opposed to the saccadic choice target) within the inactivated field, a configuration studied electrophysiologically in a categorization task²⁷, but not yet causally investigated.

Decision-related activity is probably represented broadly across the brain, and may be read out by a flexible process to support behaviour in LIP, or elsewhere^{7,18,28}. Our results call for a broader consideration of decision-making circuitry and more nuanced mechanisms for reading out decision-related activity—regardless of whether decisions are conveniently reflected, or actually computed, in the activity of a particular brain area^{23,29,30}.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 February; accepted 31 May 2016.

Published online 4 July 2016.

- Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S. & Movshon, J. A. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* **13**, 87–100 (1996).
- Shadlen, M. N. & Newsome, W. T. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* **86**, 1916–1936 (2001).
- Gu, Y., DeAngelis, G. C. & Angelaki, D. E. A functional link between area MSTd and heading perception based on vestibular signals. *Nature. Neurosci.* **10**, 1038–1047 (2007).
- Ding, L. & Gold, J. I. The basal ganglia's contributions to perceptual decision making. *Neuron* **79**, 640–649 (2013).
- Liu, S., Gu, Y., DeAngelis, G. C. & Angelaki, D. E. Choice-related activity and correlated noise in subcortical vestibular neurons. *Nature Neurosci.* **16**, 89–97 (2013).
- Hanks, T. D. *et al.* Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* **520**, 220–223 (2015).
- Pitkow, X., Liu, S., Angelaki, D. E., DeAngelis, G. C. & Pouget, A. How can single sensory neurons predict behavior? *Neuron* **87**, 411–423 (2015).
- Erich, J. C., Brunton, B. W., Duan, C. A., Hanks, T. D. & Brody, C. D. Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. *eLife* **4**, 8166 (2015).
- Cumming, B. G. & Nienborg, H. Feedforward and feedback sources of choice probability in neural population responses. *Curr. Opin. Neurobiol.* **37**, 126–132 (2016).
- Brunton, B. W., Botvinick, M. M. & Brody, C. D. Rats and humans can optimally accumulate evidence for decision-making. *Science* **340**, 95–98 (2013).
- Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
- Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* **10**, 1157–1169 (1993).
- Mazurek, M. E., Roitman, J. D., Ditterich, J. & Shadlen, M. N. A role for neural integrators in perceptual decision making. *Cereb. Cortex* **13**, 1257–1269 (2003).
- Hess, R. & Murata, K. Effects of glutamate and GABA on specific response properties of neurones in the visual cortex. *Exp. Brain Res.* **21**, 285–297 (1974).
- Newsome, W. T. & Paré, E. B. A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J. Neurosci.* **8**, 2201–2211 (1988).
- Chowdhury, S. A. & DeAngelis, G. C. Fine discrimination training alters the causal contribution of macaque area MT to depth perception. *Neuron* **60**, 367–377 (2008).
- Kiani, R., Hanks, T. D. & Shadlen, M. N. Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* **28**, 3017–3029 (2008).
- Raposo, D., Kaufman, M. T. & Churchland, A. K. A category-free neural population supports evolving demands during decision-making. *Nature Neurosci.* **17**, 1784–1792 (2014).
- Wardak, C., Olivier, E. & Duhamel, J.-R. A deficit in covert attention after parietal cortex inactivation in the monkey. *Neuron* **42**, 501–508 (2004).
- Balan, P. F. & Gottlieb, J. Functional significance of nonspatial information in monkey lateral intraparietal area. *J. Neurosci.* **29**, 8166–8176 (2009).
- Wilke, M., Kagan, I. & Andersen, R. A. Functional imaging reveals rapid reorganization of cortical activity after parietal inactivation in monkeys. *Proc. Natl Acad. Sci. USA* **109**, 8274–8279 (2012).
- Kerkhoff, G. Spatial hemineglect in humans. *Prog. Neurobiol.* **63**, 1–27 (2001).
- Li, N., Daie, K., Svoboda, K. & Druckmann, S. Robust neuronal dynamics in premotor cortex during motor planning. *Nature* **532**, 459–464 (2016).
- Crowe, D. A. *et al.* Prefrontal neurons transmit signals to parietal neurons that reflect executive control of cognition. *Nature Neurosci.* **16**, 1484–1491 (2013).
- Sarma, A., Masse, N. Y., Wang, X.-J. & Freedman, D. J. Task-specific versus generalized mnemonic representations in parietal and prefrontal cortices. *Nature Neurosci.* **19**, 143–149 (2016).
- Hanks, T. D., Ditterich, J. & Shadlen, M. N. Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nature Neurosci.* **9**, 682–689 (2006).
- Freedman, D. J. & Assad, J. A. Experience-dependent representation of visual categories in parietal cortex. *Nature* **443**, 85–88 (2006).
- Siegel, M., Buschman, T. J. & Miller, E. K. Cortical information flow during flexible sensorimotor decisions. *Science* **348**, 1352–1355 (2015).
- Heitz, R. P. & Schall, J. D. Neural chronometry and coherency across speed-accuracy demands reveal lack of homomorphism between computational and neural mechanisms of evidence accumulation. *Phil. Trans. R. Soc. B* **368**, 20130071 (2013).
- Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).

Acknowledgements We thank R. Krauzlis, C. Brody, E. Seidemann, L. Cormack, and R. Aldrich for comments on the manuscript. We thank the Brody laboratory (particularly C. Brody and J. Erlich) for inspiring the experiments, the Mauk laboratory (particularly M. Mauk, F. Riusech, and H. Halverson) for assistance with muscimol preparation, and K. Mitchell for animal support. This research was supported by the Howard Hughes Medical Institute International Student Research Fellowship to L.N.K., the McKnight Foundation grant to J.W.P., the National Eye Institute (R01-EY017366) grant to both J.W.P. and A.C.H., and the National Institutes of Health under Ruth L. Kirschstein National Research Service Awards T32DA018926 from the National Institute on Drug Abuse and T32EY021462 from the National Eye Institute.

Author Contributions L.N.K., J.L.Y. and A.C.H. designed the experiments. L.N.K. and J.L.Y. collected behavioural and electrophysiological data. L.N.K. and J.L.Y. performed pharmacological inactivations. L.N.K. analysed behavioural data. J.L.Y. analysed electrophysiological data. J.W.P. and A.C.H. guided data analysis. All authors discussed the results and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.N.K. (leor.katz@utexas.edu).

METHODS

Monkey preparation. We performed electrophysiological recordings and reversible inactivations in the MT and the LIP cortices of two rhesus macaques (subject N and subject P), female and male, aged 10 and 14 years, weighing 7.7 and 10 kg, respectively. Subject N had a custom titanium chamber that enabled access to both MT and LIP on the right hemisphere (L9, P2), guided by MRI. Subject P had a cilux chamber (Crist Instruments) over the right LIP (L12, P5) and another over the left V1 for a posterior approach to MT (L17, P17). Standard surgical procedures were applied³¹. All experimental protocols were approved by The University of Texas Institutional Animal Care and Use Committee and in accordance with National Institute of Health standards for care and use of laboratory animals.

The subject sat comfortably while head-posted in a primate chair (Crist Instruments), facing a linearized 55 inch LCD (LG) monitor (resolution = 1,920 × 1,080 pixels, refresh rate = 60 Hz, background luminance = 26.49 cd m⁻²) at a distance of 118 cm, in a dark room. Eye position was recorded using an Eyelink 1000 eye tracker (SR Research), sampled at 1 kHz. A solenoid-operated reward system was used to deliver liquid reward to the monkey. Stimuli were generated by using the Psychophysics Toolbox³² in MATLAB (MathWorks), and task events and neural responses were recorded (Plexon) using a Datapixx I/O box (Vpixon) for precise temporal registration. All of these systems were integrated using the PLDAPS system³³.

General procedure and experimental design. Recording sessions in either MT or LIP began by lowering an electrode to the known location of the area based on previous mapping and recording sessions. Anatomical identification (MR guided in monkey N; previously established in monkey P³¹) was followed by functional identification (mapping receptive/response fields of MT and LIP neurons, detailed below). Inactivations of either area began by lowering both a cannula and multi-channel electrode array to the region of interest, collaterally, at least 1 mm apart. The electrode array was used to (i) confirm that the cannula was within the target cortex, (ii) to record electrophysiological responses to relevant task events pre-infusion, and (iii) to confirm the electrophysiological silencing of neurons during and after the infusion. Thus, while it is not feasible to precisely measure the inactivated proportion of an area, we do confirm the silencing of a large swath (approximately 2.5 mm in radius), on every session (detailed in infusion protocol section, below).

MT inactivation was predicted to disrupt motion direction discrimination sensitivity within a specific region of contralateral space, consistent with MT retinotopic organization^{15,16}. The behavioural consequence of MT inactivation was measured by comparing psychophysical performance in the direction-discrimination task, before and after muscimol infusion, within the same experimental session, with the motion stimulus placed inside the inactivated region of space. LIP inactivation was predicted to disrupt spatial selection to contralateral space more generally^{8,19–21,34,35}, noting that LIP receptive fields are large and that the topographic organization is less precise than in earlier visual areas³⁶. The behavioural consequence of LIP inactivation was measured by comparing the proportion of contralateral choices in a double-target memory-guided free-choice task, before and after muscimol infusion, within the same session. To measure the effect of LIP inactivation in the direction-discrimination task, we compared psychophysical performance between pairs of sessions, baseline and treatment, in which the treatment session was a muscimol, saline, or sham infusion treatment. The paired sessions typically took place one day apart at the same time of day, and after a similar number of tasks and trials, to minimize the impact of within-session fatigue or motivation on behaviour. Behavioural data were collected 15–30 min after muscimol/saline/sham infusion end, and were always completed within 150 min. An additional 16 control pairs (without saline or sham manipulation) were collected to better estimate session-to-session variability. Statistical results do not depend on the inclusion/exclusion of these control session pairs. The number of paired sessions (Extended Data Fig. 6c and Extended Data Table 2) was chosen to match or exceed that in similar inactivation studies. The experiments were not randomized, and the investigators were not blinded to the experimental conditions.

Direction discrimination task. The principal task was a motion direction discrimination task. Subjects were required to discriminate the net direction of a motion stimulus and communicate their decision with an eye movement to one of two targets. The sequence of task events is presented in Fig. 1a. The timing of each event was randomly jittered from trial to trial (Fig. 1b). A trial began with the appearance of a fixation point. Once the monkey acquired fixation and held for 400–1,200 ms (uniform distribution), two targets appeared and remained visible until the end of the trial. 200–1,000 ms after target onset, the motion stimulus was presented at an eccentricity of 5–7° for 1,050 ms. The fixation point was extinguished 200–1,000 ms after motion offset, and the subject was required to shift

its gaze towards one of the two targets within 600 ms (saccade end points within 3° of the target location were accepted).

We used a reverse-correlation motion stimulus inspired by the classic moving-dot stimulus¹⁵ in which motion was in either one direction or the opposite, with varying motion strength. The motion stimulus consisted of 19 non-overlapping Gabor elements arranged in a hexagonal grid (5–7° across, scaled by eccentricity). The individual elements were set to approximate the receptive field size of a V1 neuron and the entire motion stimulus approximated the receptive field size of an MT neuron. Motion was presented by varying the phase of the sine-wave carrier of the Gabors. Each Gabor underwent a sinusoidal contrast modulation with independent random phase to prevent perceptual ‘pop-out’ of individual drifting elements. Gabor spatial frequency (0.9 cycles per °, $\sigma = 0.1 \times \text{eccentricity}$) and temporal frequency (7 Hz for monkey N, 5 Hz for monkey P, yielding velocities of 7.77 and 5.55° per s, respectively) were selected to match the approximate sensitivity of MT neurons.

Each trial comprised seven consecutive motion pulses lasting 150 ms each (9 video frames), producing a pulse sequence of 1,050 ms in duration. On any given pulse X_i , a number of Gabors would have their carrier sine waves drift in unison to produce motion (‘signal’ Gabors), and the remaining would counter-phase flicker (‘noise’ Gabors). Signal Gabors on pulse X_i were assigned at random within the grid and all signal Gabors drifted in the same direction.

Motion strength was defined as the proportion of signal Gabors out of the total, the value of which was drawn from a Gaussian distribution, $X_i \sim N(\mu_k, \sigma)$ and rounded to the nearest integer, where μ_k was set to one of five values at random: -50%, -12%, 0%, 12%, and 50% (negative sign indicates motion in the opposite direction), and σ was set to 15%. Thus, although each pulse within a sequence could take on any value (or sign) from distribution $N(\mu_k, \sigma)$, the expectation of a sequence would be μ_k . Motion strength was then z scored over all sessions for each monkey separately.

On the motion strength axis, we use positive values to indicate motion towards the hemifield contralateral to the LIP under study, and negative values to indicate motion towards the hemifield ipsilateral to the LIP under study. We use the term ‘proportion choices’ to refer to the proportion of choices towards the contralateral target. For consistency, we maintain this convention throughout the paper, such that even on MT inactivation sessions, psychometric performance is evaluated in relation to the LIP under study.

The monkey was rewarded for selecting the target consistent with the sign of the motion pulse sequence sum (the net direction), independent of the distribution μ_k from which they were drawn. On trials that summed to exactly zero, the monkey was rewarded at random. 10% of trials consisted of a frozen random seed, generating identical pulse sequences. In addition to the direction discrimination task described here, we performed a subset of experiments ($n = 2$) using the classical moving dots stimulus¹⁵ with motion coherence values of 0, 3.2, 6.4, 12.8, 25.6 and 51.2% (Extended Data Fig. 6c).

Free-choice task. A free-choice task was used to measure spatial bias to one target over another and confirm a behavioural consequence of LIP inactivation^{8,21,35}. The task was performed before and after every LIP inactivation ($n = 21$ during experiments using the standard direction discrimination task, $n = 13$ during other experiments, see for example, Extended Data Fig. 6c). The sequence of events within the free-choice task is illustrated in Fig. 3a, b. Trials began with the appearance of a central fixation point. At a random time after acquiring fixation (500–900 ms), two targets were simultaneously flashed for a brief 200 ms. Subjects were required to maintain fixation until the fixation point disappeared (600 to 3,000 ms after target flash), and then saccade to either of the remembered locations of the two targets. On every trial, target position was determined independently from one another and at random, drawn from a 2D Gaussian with a mean of either (-12, 0)° (left target) or (12, 0)° (right), and a standard deviation of 2–4° for x and 3–5° for y position. Means and standard deviations were sometimes adjusted online to better position the distributions within the LIP receptive field (when recorded) or LIP inactivated field (when inactivated).

A trial was successfully completed when the monkey’s saccade entered a circular window (unobservable to the monkey) around either target and held for 300–500 ms (window radius scaled by $0.35^\circ \times \text{eccentricity}$, minimum: 3°). Successfully completed free-choices were rewarded on 70% of trials irrespective of the target chosen for monkey N, and 100% of trials for monkey P. Monkey N also performed memory-guided saccades to single targets (30% of trials, randomly interleaved) that appeared randomly in space (uniform distribution), and were rewarded 100% of the time. The adjustments in subject N’s task were performed to prevent a spatial bias and encourage exploration. Overall performance and inactivation effects were similar between monkeys despite differences in task parameters. **Behavioural analysis.** All analyses were performed in MATLAB (MathWorks). Responses in the direction discrimination task were analysed with a maximum

likelihood fit of a two parameter logistic function³⁷ assuming a Bernoulli distribution of binary choices, in which the probability of a contralateral choice is P and ipsilateral choice is $1 - P$, where P is given by:

$$P = \frac{1}{1 + e^{-\beta(x-\alpha)}}$$

where x is the motion strength value (z -scored), α is the bias parameter (reflecting the midpoint of the function in units of motion strength), and β is the slope (that is, sensitivity, in units of log-odds per motion strength). Error estimates on the parameters were obtained from the diagonal of the inverse Hessian (second derivative matrix) of the negative log-likelihood. A four-parameter model including sub-perfect response rates for the top and bottom asymptotes⁸ was also considered, but did not confer any advantage over the two-parameter model nor change analysis results, and so we focused on the simpler two-parameter fit (Extended Data Table 1). The first 10–30 trials of every session were ‘instruction’ (or warm-up) trials (in which motion strength was set to maximal) and were excluded from analysis. Median session length for all baseline and treatment sessions was 409 trials. Sessions were excluded from analysis if the animal either completed less than 250 trials or performed poorly (lapse rate >10%). For inactivation sessions, all sessions were included regardless of performance. A single inactivation session in monkey P was aborted due to a leak in the infusion system, and was not included in the analysis.

Animal strategy in the direction discrimination task (Fig. 2f, g) was measured by computing psychophysical weights via logistic regression, where the probability of the binary choice $Y \in \{0, 1\}$ on every trial is given by

$$p(Y|w, X) = e^{YXw} / (1 + e^{Xw})$$

where X is a matrix of the seven pulse values on each trial, augmented by a column of ones to capture the bias term, and w is a vector of the monkey's weights. We computed the maximum likelihood estimate of the weight vector w using MATLAB's glmfit function.

In the free-choice task, spatial bias was computed as the proportion of choices to the target contralateral to the LIP under study. Saccade onset and offset were detected in every task by identifying the time at which eye velocity exceeded 30° per s (onset) and returned below 50° per s (offset). We only analysed saccades on trials where the task was completed successfully (no broken fixations and no saccades outside of the target windows). Saccades were analysed for reaction time, amplitude, duration, and error amplitude (distance of saccadic end point from saccadic target). Saccadic reaction times less than 100 ms from the go signal were excluded to ensure that only task relevant saccades are analysed.

Neuronal recordings. Recordings were performed in areas MT and LIP with either single-channel glass coated tungsten electrodes (Alpha Omega) or multi-electrode arrays (Plexon U or V Probe). Neuronal signals were amplified, band-pass filtered, digitized, and saved (Plexon MAP server). Neural waveforms passing a manually-set threshold were isolated for online mapping of their receptive fields (both MT and LIP) and directional tuning (MT).

MT receptive field locations were hand mapped using drifting dot stimuli in a circular aperture. Once the retinotopic location was identified, direction preference and selectivity were measured using drifting dot stimuli at 100% coherence in 12 directions. LIP receptive field locations were mapped with a memory-guided delayed saccade task³⁸.

In monkey P, offline spike sorting was performed by hand refinement of a standard clustering algorithm (Plexon Offline Sorter v3). Single unit isolation quality was established using SNR³⁹. In monkey N, spike sorting was performed by fitting a mixture of Gaussians model to clipped waveforms in a reduced dimensional space⁴⁰. In both monkeys, sorting was refined by maximum *a posteriori* estimation of a model, where the multi-electrode voltage was the linear superposition of Gaussian white noise and the spike waveforms^{41,42}.

Neuronal analysis. Peri-stimulus time histograms (PSTHs) were computed by aligning spike times to events (motion onset or saccade time), binned at 10 ms resolution, and smoothed with a Gaussian kernel with standard deviation of 25 ms. Trial motion strengths were binned into three groups: between 0 and 0.25, between 0.25 and 1, and greater than 1. We averaged spike rates separately for the three motion strengths for each choice. Note that these motion strengths correspond to a narrower range than that used in previous studies¹¹, selected to encourage longer integration times. This is evident in the PSTHs (narrow dynamic range) and psychometric functions (fewer data points in the asymptotic range of behaviour).

Choice probability. Choice probability is a metric used to measure the predictive relationship between neural responses and choice, independent of stimulus strength. It is defined as the area under the receiver operating characteristic curve (ROC) for a pair of spiking response distributions sorted by choice^{1,43}.

We quantified choice probability using trials that had zero expected motion and were repeated with identical random seeds (that is, had no stimulus variation, ‘frozen noise’). Sometimes more than one random seed was repeated in a session, in which case we calculated the spiking response distributions for each seed separately, subtracted the mean, and then combined them, similar to an analysis known as grand choice probability¹. Neurons with >20 frozen repeats were included (90/94 MT cells, 96/113 LIP cells), and significance testing against the null (that is, choice probability = 0.5) was performed using a Student's *t*-test. In MT, we counted spikes during the motion epoch (1,050 ms). In LIP, we counted spikes over a 400 ms window counting backwards from the 100 ms before the saccade.

Infusion protocol. Infusions were performed by lowering an infusion cannula into grid locations that had previously yielded the largest number of selective cells during the recording phase of the study (Extended Data Fig. 1). The cannula (31–32 gauge) was lowered alongside a multi-electrode array, at least 1 mm away (Fig. 2a). The two were lowered to target cortical areas where functional identification took place (mapping). Infusion was then performed, and electrophysiological silencing was confirmed on the recording electrodes, typically within 15 min of infusion start.

Infusions were performed with a syringe pump (Harvard Apparatus) through a single and direct line to the cannula (constant rate of 0.1–0.4 $\mu\text{L min}^{-1}$, 15–30 min), in agreement with infusion parameters proposed in ref. 44. We delivered 6.66–8 $\mu\text{g } \mu\text{L}^{-1}$ muscimol (in phosphate buffered saline) at volumes of 5–12 μL (mean 7.4 μL), netting a total mass of 40–80 μg (mean 56.4 μg). This protocol was chosen to match the very high end of ranges used previously in order to maximize neural inactivation. Infusions were typically made at multiple depths within a single cannula track. On 5 of the 21 main LIP inactivation sessions, more than one cannula was lowered (Extended Data Table 2). Cannulae were left *in situ* for at least 15 min after infusion end. Saline infusions followed the same protocol and included both a cannula and multi-electrode array. Sham infusions included only a multi-electrode array but followed similar timings, including the operation of the syringe pump with no syringe attached.

Spatial and temporal extent of inactivation. Previous analyses of the spatial extent of muscimol inactivation have estimated the functional silencing to cover a spherical radius of roughly 2–3 mm^{34,45–47}. The study most comparable to ours, by Liu *et al.*³⁴, co-infused muscimol and Manganese (Mn) into LIP of awake macaques and imaged the spread. They also estimated a cortical silencing of approximately 2–3 mm in radius, consistent with the linear dependence of volume distribution (mm^3) on infusion volume (μL)⁴⁸.

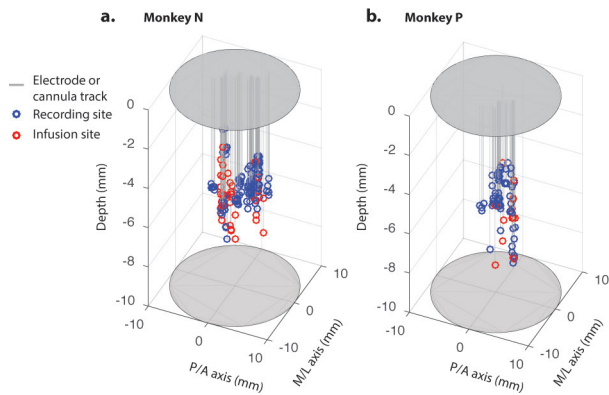
In our experiments, lowering both a multi-electrode array and infusion cannula collaterally (Fig. 2b) enabled direct confirmation of neural silencing at known distances from the cannula tip. This places a lower bound on the spatial extent of functional inactivation. Although our standard protocol placed the multi-electrode array 1 mm away from the cannula tip, we sometimes lowered a second array, 2 or 3 mm away. On these sessions too, we observed silencing on most recording channels. Taken together, we conservatively estimate neural inactivation in LIP to span a radius of at least 2.5 mm, silencing large swaths of LIP, while primarily targeting its ventral portion^{34,49}. For inactivations of this spatial magnitude, there is no evidence that larger inactivations result in larger behavioural deficits²⁰. Similarly, we did not observe a dose–response function in our own data (Extended Data Fig. 3c–e).

On a few occasions, residual firing persisted despite near-complete silencing of electrophysiological activity (example shown in Fig. 2b, voltage traces, channels 5 and 6). We tested the selectivity of residual firing with the appropriate mapping task (motion for MT, memory guided saccades for LIP) and found that these spikes did not respond selectively, indicating that these residual spikes likely emanate from afferent fibres terminating within the inactivated area⁵⁰.

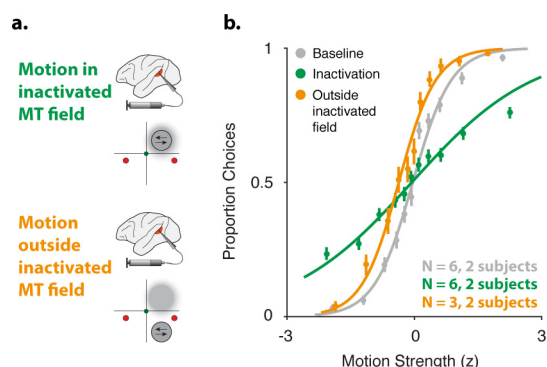
Previous LIP inactivation studies found no evidence to support within-session compensation that manifests behaviourally^{19,20,34,47,51}, but see ref. 21. Studies that report the temporal effect of LIP inactivation find an increase in the impact over time, not a decrease^{19,51}. Regardless, we explored the time course of psychophysical performance within a session (Extended Data Fig. 4), and also evaluated compensation on longer time scales, across sessions, to explore the possibility of increasing behavioural robustness to inactivation that might develop over time (Extended Data Fig. 5).

31. Meister, M. L. R., Hennig, J. A. & Huk, A. C. Signal multiplexing and single-neuron computations in lateral intraparietal area during decision-making. *J. Neurosci.* **33**, 2254–2267 (2013).
32. Brainard, D. H. The Psychophysics Toolbox. *Spat. Vis.* **10**, 433–436 (1997).
33. Eastman, K. M. & Huk, A. C. PLDAPS: a hardware architecture and software toolbox for neurophysiology requiring complex visual stimuli and online behavioral control. *Front. Neuroinform.* **6**, 1 (2012).
34. Liu, Y., Yttri, E. A. & Snyder, L. H. Intention and attention: different functional roles for LIPd and LIPv. *Nature Neurosci.* **13**, 495–500 (2010).

35. Zirnsak, M., Chen, X., Lomber, S. G. & Moore, T. Effects of reversible inactivation of parietal cortex on the processing of visual salience in the frontal eye field. *Proc. Conference Soc. Neurosci.* (2015).
36. Patel, G. H. *et al.* Topographic organization of macaque area LIP. *Proc. Natl Acad. Sci. USA* **107**, 4728–4733 (2010).
37. Wichmann, F. A. & Hill, N. J. The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept. Psychophys.* **63**, 1293–1313 (2001).
38. Gnadt, J. W. & Andersen, R. A. Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* **70**, 216–220 (1988).
39. Kelly, R. C. *et al.* Comparison of recordings from microelectrode arrays and single electrodes in the visual cortex. *J. Neurosci.* **27**, 261–264 (2007).
40. Tolias, A. S. *et al.* Recording chronically from the same neurons in awake, behaving primates. *J. Neurophysiol.* **98**, 3780–3790 (2007).
41. Pillow, J. W. *et al.* Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* **454**, 995–999 (2008).
42. Pillow, J. W., Shlens, J., Chichilnisky, E. J. & Simoncelli, E. P. A model-based spike sorting algorithm for removing correlation artifacts in multi-neuron recordings. *PLoS One* **8**, e62123 (2013).
43. Celebriini, S. & Newsome, W. T. Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *J. Neurosci.* **14**, 4109–4124 (1994).
44. Noudoost, B. & Moore, T. A reliable microinjectrode system for use in behaving monkeys. *J. Neurosci. Methods* **194**, 218–223 (2011).
45. Martin, J. H. Autoradiographic estimation of the extent of reversible inactivation produced by microinjection of lidocaine and muscimol in the rat. *Neurosci. Lett.* **127**, 160–164 (1991).
46. Arikan, R. *et al.* A method to measure the effective spread of focally injected muscimol into the central nervous system with electrophysiology and light microscopy. *J. Neurosci. Methods* **118**, 51–57 (2002).
47. Yttri, E. A., Wang, C., Liu, Y. & Snyder, L. H. The parietal reach region is limb specific and not involved in eye-hand coordination. *J. Neurophysiol.* **111**, 520–532 (2014).
48. Heiss, J. D., Walbridge, S., Asthagiri, A. R. & Lonser, R. R. Image-guided convection-enhanced delivery of muscimol to the primate brain. *J. Neurosurg.* **112**, 790–795 (2010).
49. Lewis, J. W. & Van Essen, D. C. Mapping of architectonic subdivisions in the macaque monkey, with emphasis on parieto-occipital cortex. *J. Comp. Neurol.* **428**, 79–111 (2000).
50. Chapman, B., Zahs, K. R. & Stryker, M. P. Relation of cortical cell orientation selectivity to alignment of receptive fields of the geniculocortical afferents that arborize within a single orientation column in ferret visual cortex. *J. Neurosci.* **11**, 1347–1358 (1991).
51. Kubanek, J., Li, J. M. & Snyder, L. H. Motor role of parietal cortex in a monkey model of hemispatial neglect. *Proc. Natl Acad. Sci. USA* **112**, E2067–E2072 (2015).

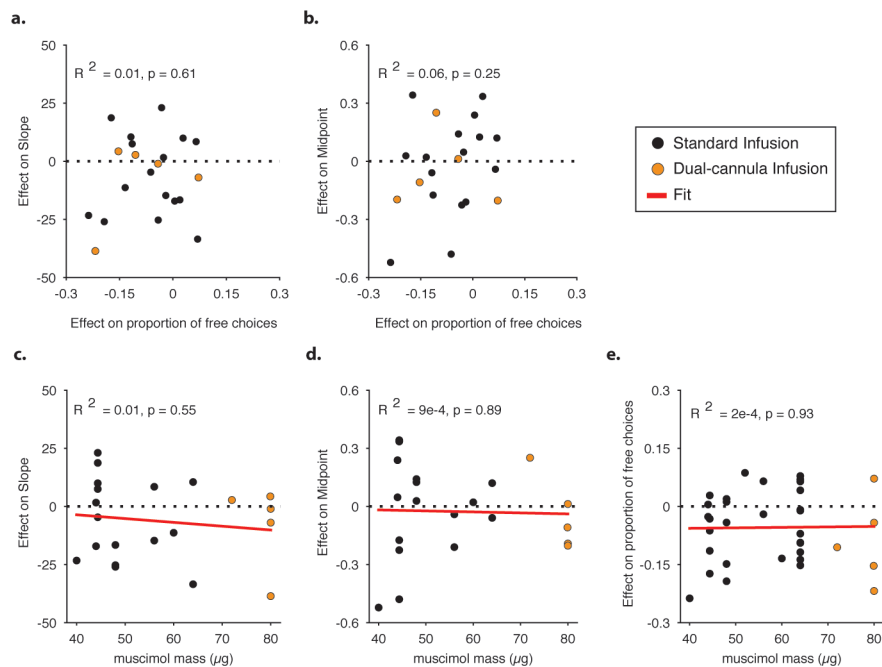


Extended Data Figure 1 | Location of LIP recording and muscimol infusion sites. a, b, The recording (blue circles) and infusion sites (red) for monkey N (**a**) and monkey P (**b**) along the medial-lateral (M/L) and posterior-anterior (P/A) axes within the chamber (demarcated by the ovals). Electrode and cannula tracks are represented by the grey lines (with a small jitter on the x - y plane for better visualization). The mean infusion depths were 7.12 ± 1.15 mm (monkey N) and 7.03 ± 1.39 mm (monkey P) (the microdrive was zeroed below dura mater and just above the cortical surface). Given the estimated spread of muscimol described in the main text, the inactivations targeted a substantial territory of the ventral portion of LIP⁴⁹. Even though a functional distinction with depth has been proposed³⁴, we emphasize that the critical component of our protocol was targeting the precise locations at which we measured canonical decision-related activity in LIP.



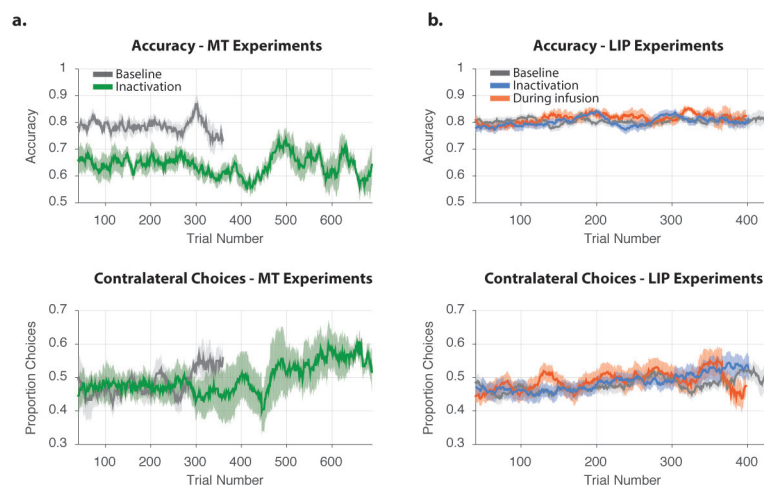
Extended Data Figure 2 | Direction discrimination sensitivity is restored when motion is placed outside of the inactivated MT field.

a. Illustration of MT inactivation along with the estimated inactivated field (grey cloud), for two experimental geometries: motion stimulus placed inside the inactivated MT field (top) and motion placed outside the inactivated MT field (bottom). **b.** Average psychophysical data for baseline and muscimol treatment pairs (grey and green, respectively, same data as Fig. 2c, $n = 6$; 3 in monkey N; 3 in monkey P) and psychophysical data collected during muscimol treatment, with the motion stimulus outside of the inactivated MT field (orange, $n = 3$). Direction discrimination sensitivity is restored to baseline levels in these sessions. Error bars on points show ± 1 s.e.m. over all trials.



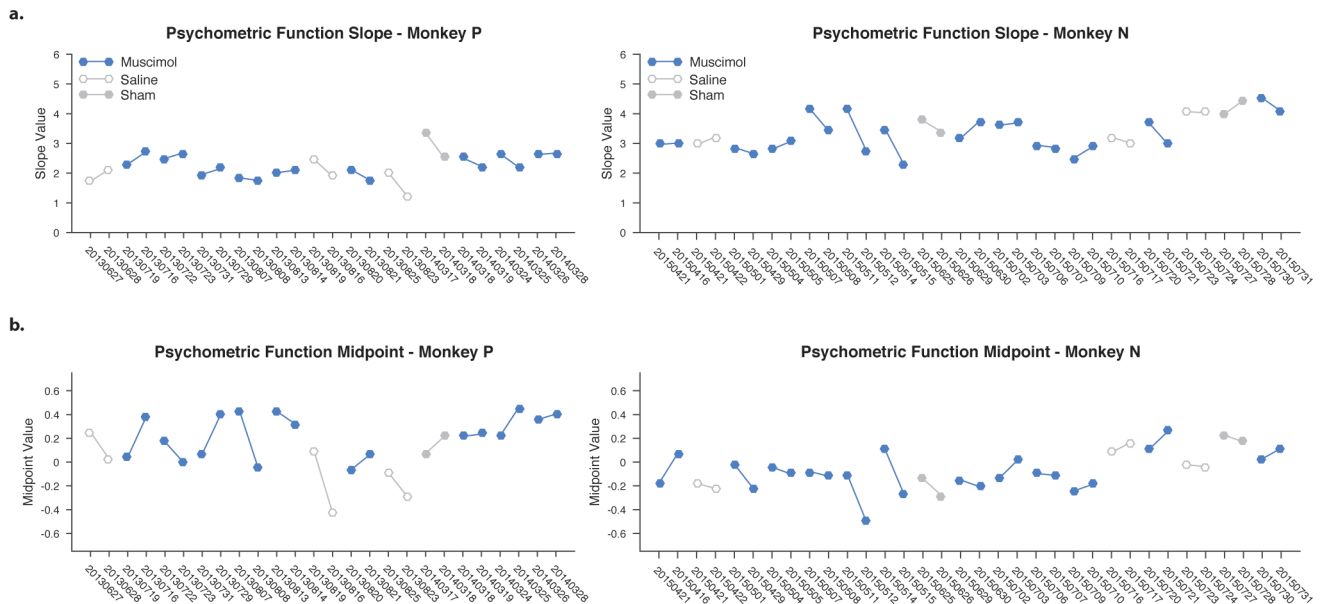
Extended Data Figure 3 | No relationship between effect magnitude in control task, effect magnitude in direction discrimination task, and muscimol mass. **a, b,** The relationship between the effect of LIP inactivation in the free-choice task (that is, shift in proportion of contralateral choices from baseline to muscimol treatment) and the effect of LIP inactivation in the direction discrimination task on sensitivity (percentage change in psychometric function slope, **a**) and bias (shift in normalized motion strength, **b**). R^2 and associated P values of a Pearson correlation are indicated on individual plots ($n = 21$; 12 in monkey N; 9 in monkey P). Orange data points indicate sessions in which muscimol

was infused from two cannulae simultaneously into LIP. **c–e,** Dose–response functions between muscimol mass and the effect in the direction discrimination task on slope (**c**, same units as **a**), bias (**d**, same units as **b**), and the effect in the free-choice task (**e**, same units as **a, b**). For **e**, we used free-choice sessions that took place on the same days as the direction discrimination task ($n = 21$) along with an additional 13 sessions that took place during other inactivation experiments under similar conditions ($n = 34$ in total; 14 in monkey N; 20 in monkey P; as in Fig. 3d). R^2 , associated P values and regression lines are indicated on the plots (linear regression).



Extended Data Figure 4 | Time course of accuracy and bias within sessions. Accuracy and bias in the direction discrimination task were computed over time by taking a running mean of correct and contralateral choices, respectively (sliding window of 40 trials). **a**, Inactivation in area MT ($n = 6$, green curve; 3 in monkey N; 3 in monkey P) had a clear and consistent impact on behavioural accuracy compared to baseline ($n = 6$, grey), but did not have systematic effects on bias (bottom), consistent with our results from the fitted psychometric functions (main text). Panels show data from trial 40 (sliding window size) to the median trial length of each group of experiments (variable session lengths contribute to increased variability at later trials). Error bars show ± 1 s.e.m. between sessions. **b**, Inactivations in area LIP ($n = 21$, blue curve; 12 in monkey N;

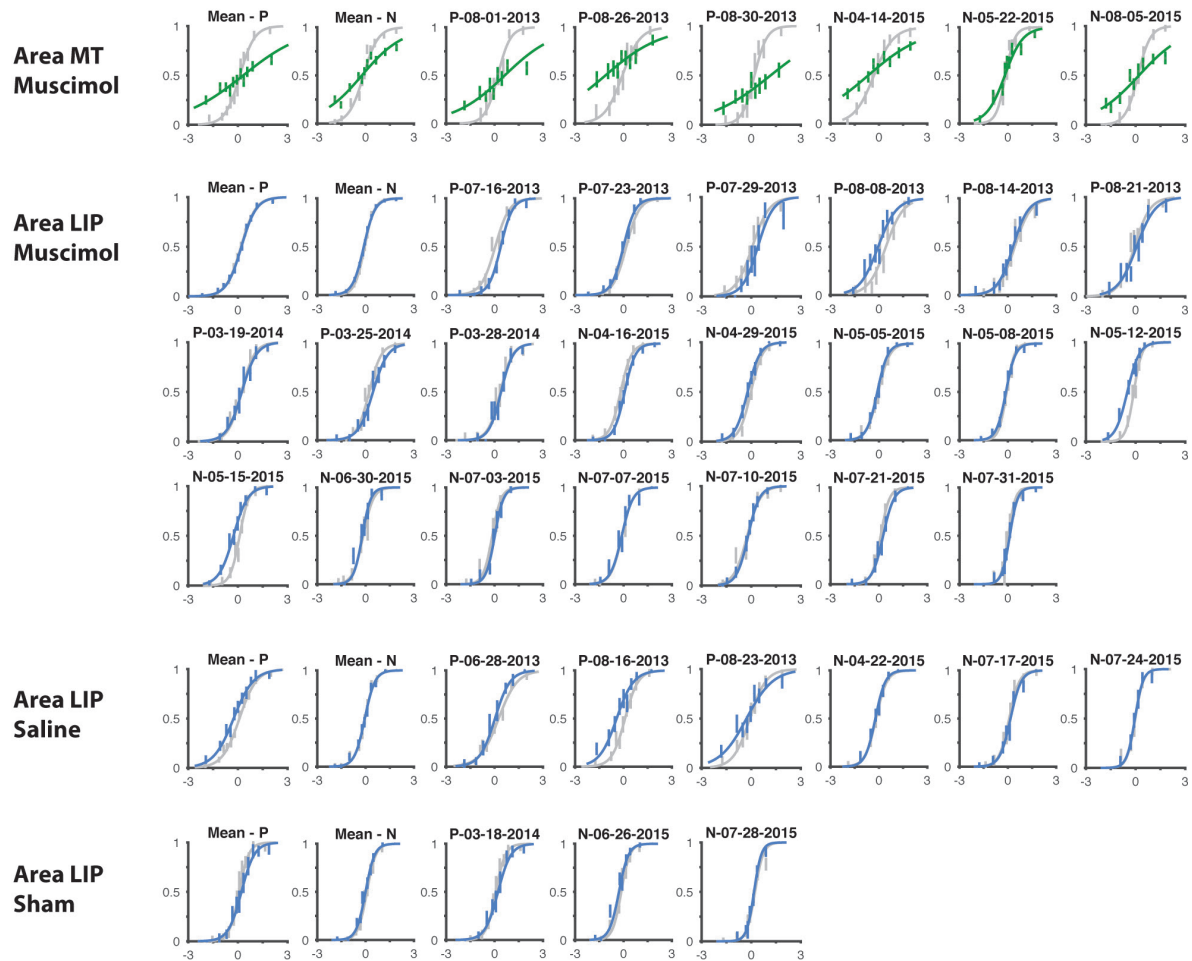
9 in monkey P) yielded no systematic trends in either accuracy (top) or bias (bottom) compared to baseline ($n = 21$, grey), indicating that within-session compensation is unlikely. Panel format same as in **a**. We also investigated whether compensation may have taken place before we began collecting the 'inactivation' data set, or during the first 10–30 instruction (warm-up) trials. On 13 of the 21 LIP inactivation sessions, we collected a third data set (in addition to the standard paired baseline and inactivation data sets), in which psychophysical performance was monitored during the time muscimol was being infused (during infusion, orange curve). No systematic changes in accuracy or bias were observed in this exploratory data set either, further arguing against compensation on the time scales of our manipulations and measurements.



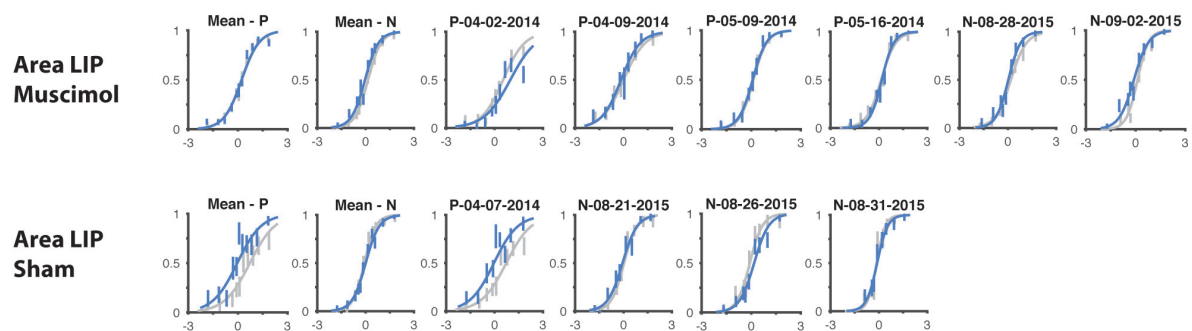
Extended Data Figure 5 | Psychophysical performance in the direction discrimination task across sessions. Panels show data from monkey P (left) and monkey N (right), for all baseline and treatment pairs: muscimol (blue, $n = 21$), saline (unfilled grey, $n = 6$) and sham (filled grey, $n = 3$). Each pair consists of two sessions that took place in close succession (typically on consecutive days), at a similar time of day, after a similar number of preceding tasks and trials, and is represented by two markers connected by a line. Additional control pairs with no saline/sham manipulation ($n = 16$) are not presented, for visual clarity. **a**, Psychometric function slope over sessions. No significant change in slope was present over time, evaluated by linear regression, for either monkey P ($P = 0.22$) or N ($P = 0.63$). When considering the difference in slope between baseline and treatment pairs, monkey P exhibited a small decrease

(regression line slope = -0.07 , $P = 0.023$). However, a similar effect was seen in the interleaved controls (saline and sham, grey markers), indicating that this pattern likely reflects nonspecific trends in performance across back-to-back pairs of experiments. Monkey N had no significant change ($P = 0.92$). **b**, Psychometric function midpoint over sessions. No significant change was observed in the session-to-session midpoint values, evaluated by linear regression, for either monkey P ($P = 0.44$) or monkey N ($P = 0.24$). When considering the difference in midpoint value for each data set pair over time (that is, muscimol treatment – baseline), no significant change was detected either ($P = 0.98$ and $P = 0.4$ for monkey P and N, respectively). The x axis dates are in the year, month, date, yyyyymmdd format.

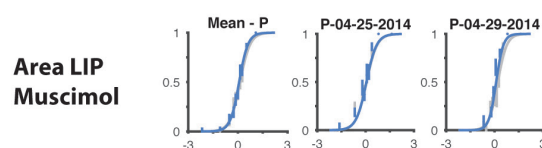
a. Standard Geometry



b. Both targets in inactivated field



c. Newsome dots



Extended Data Figure 6 | Psychophysical performance for all individual baseline and treatment session pairs. a–c. All pairs of baseline and treatment sessions for all treatment types: muscimol, saline, and sham, (control pairs with no saline/sham manipulation are similar but not presented, for visual clarity) for all variants of the direction discrimination task: standard geometry (a), both targets in inactivated field (b), and Newsome dots (c), for both LIP and MT inactivation. In all panels, the

abscissa represents motion strength towards the direction contralateral to the LIP under study, the ordinate represents the proportion of contralateral choices. The grey curve is baseline, and the coloured curve is treatment. The first panel in each section presents mean psychophysical performance for each monkey over sessions. Subsequent panels present individual session pairs. Error bars are s.e.m. over all trials.

Extended Data Table 1 | Parametric and nonparametric analysis of psychophysical data, for two- and four-parameter psychometric functions

Statistical			Muscimol Infusions				Control Infusions				Muscimol vs. Control Infusions				
test	model	monkey	midpoint	slope	minLapse	maxLapse	midpoint	slope	minLapse	maxLapse	midpoint	slope	minLapse	maxLapse	
Student's t	pmf2	N	0.542	0.1734			0.986	0.2444			0.149	0.2367			
		P	0.6731	0.7982			0.2353	0.2166			0.0208	0.2461			
		Both	0.9306	0.1659			0.3693	0.1092			0.2606	0.0704			
	pmf4	N	0.4028	0.0243	0.7585	0.2352	0.8213	0.2827	0.6065	0.6228	0.8213	0.2827	0.6065	0.6228	
		P	0.6747	0.459	0.3554	0.8337	0.1377	0.3261	0.5675	0.3111	0.1377	0.3261	0.5675	0.3111	
		Both	0.8163	0.0552	0.3904	0.9091	0.2169	0.1595	0.7461	0.2734	0.2169	0.1595	0.7461	0.2734	
	WSRST	pmf2	N	0.791	0.3394			0.9308	0.2305			0.184	0.1496		
			P	0.5703	0.9102			na	na			0.0503	0.3301		
			Both	0.9032	0.3219			0.4432	0.1036			0.3898	0.0324		
pmf4		N	0.6221	0.021	0.5186	0.2334	0.9032	0.2305	0.3754	0.4761	0.184	0.4887	0.184	0.5125	
		P	0.4961	0.4258	0.8203	0.8203	na	na	0.875	0.625	0.0503	0.6042	0.8252	0.8252	
		Both	0.8213	0.0325	0.414	0.566	0.3533	0.1353	0.2758	0.3674	0.3543	0.0895	0.4942	0.2009	

The entries show *P* values for two types of statistical analyses: the parametric Student's *t*-test and the non-parametric Wilcoxon signed-rank sum test (WSRST). The tests were performed on model parameters fit to individual sessions. We present data for the standard two-parameter psychometric function (pmf2), and for an exploratory four-parameter psychometric function (pmf4). Muscimol infusions: paired tests compared muscimol baseline sessions to muscimol treatment sessions. Control infusions: paired tests compared saline/sham/control baseline sessions to saline/sham/control treatment sessions. Muscimol versus control infusion: unpaired tests compared muscimol treatment sessions to saline/sham/control treatment sessions. na, not enough data.

Extended Data Table 2 | Infusion details for all treatment sessions

Task	Area	Monkey	Date	Treatment	Cannula Tracks (#)	Positioning Grid (x, y)	Infusion sites within Track (#)	Average depth (mm)	Total volume (μl)	Total mass (μg)
Standard task geometry	MT	P	20130801	Muscimol	1	(2, -1)	2	8	5	33.3
			20130826	Muscimol	1	(2, 0)	2	10	9.7	64.3
			20130830	Muscimol	1	(2, 0)	2	11.2	8.5	56.6
		N	20150414	Muscimol	1	(5, -4)	3	10.7	4	32
			20150522	Muscimol	1	(5, -4)	2	10.5	5	40
			20150805	Muscimol	1	(4, -4)	2	6.9	5	40
	LIP	P	20130628	Saline	1	(2, -1)	1	7	6.7	-
			20130716	Muscimol	1	(3, 0)	1	6.5	6.7	44.4
			20130723	Muscimol	1	(3, 0)	1	7	6.7	44.4
			20130729	Muscimol	1	(2, -1)	1	7	6.7	44.4
			20130808	Muscimol	1	(3, 0)	1	6	6.7	44.4
			20130814	Muscimol	1	(3, 0)	1	7	6.7	44.4
			20130816	Saline	1	(3, 0)	1	6	6.7	-
			20130821	Muscimol	2	(3, 0); (0, 3)	2; 2	7; 7	12	79.9
			20130823	Saline	1	(3, 0)	1	7	6.7	-
			20140318	Sham	-	-	-	-	-	-
			20140319	Muscimol	1	(3, 0)	1	7	5	40
			20140325	Muscimol	1	(3, 0)	1	7	6	48
			20140328	Muscimol	2	(3, 0); (1, -3)	2; 2	7.5; 7.5	10	80
		N	20150416	Muscimol	1	(2, 4)	2	6.3	6	48
			20150422	Saline	1	(2, 4)	1	7.6	5	-
			20150429	Muscimol	1	(-2, 3)	2	6.6	7.5	60
			20150505	Muscimol	1	(-2, 3)	2	6.5	5.5	44
			20150508	Muscimol	1	(-1, 3)	2	8.6	5.5	44
			20150512	Muscimol	2	(-2, 3); (3, 4)	2; 2	7.9; 7.9	9	72
			20150515	Muscimol	1	(-2, 3)	3	7.6	7	56
			20150626	Sham	-	-	-	-	-	-
			20150630	Muscimol	1	(-2, 3)	2	8	7	56
			20150703	Muscimol	1	(-3, 2)	3	7.1	8	64
			20150707	Muscimol	2	(-3, 2); (2, 4)	3; 3	7.9; 7.9	10	80
			20150710	Muscimol	2	(-3, 2); (2, 3)	2; 2	6.2; 6	10	80
			20150717	Saline	1	(-3, 2)	2	5.3	6	-
			20150721	Muscimol	1	(-3, 2)	2	6.1	8	64
			20150724	Saline	1	(-3, 2)	2	6	5	-
			20150728	Sham	-	-	-	-	-	-
			20150731	Muscimol	1	(-3, 2)	3	6	6	48
Both targets in inactivated field	LIP	P	20140402	Muscimol	1	(3, 0)	2	7.5	6	48
			20140407	Sham	-	-	-	-	-	-
			20140409	Muscimol	1	(3, 0)	4	7	8	64
			20140509	Muscimol	1	(3, 0)	3	9	8	64
			20140516	Muscimol	1	(3, 0)	3	9	8	64
		N	20150821	Sham	-	-	-	-	-	-
			20150826	Sham	-	-	-	-	-	-
			20150828	Muscimol	1	(-3, 2)	3	7	6	48
			20150831	Sham	-	-	-	-	-	-
			20150902	Muscimol	1	(-3, 2)	4	6.2	6.5	52
Newsome Dots	LIP	P	20140425	Muscimol	1	(3, 0)	3	7.2	8	64
			20140429	Muscimol	1	(3, 0)	3	7.6	8	64

The table presents all infusion sessions run over the course of the study for all infusion types (muscimol, saline, sham), in either MT or LIP. Infusions are sorted by date within each task, for each monkey separately. Positioning grid values are relative to chamber centres (see Methods for stereotactic coordinates). Average depth refers to the average depth across all infusion sites within a given cannula track. Total volume and total mass refer to the sum over all infusion sites and tracks.

Resolving early mesoderm diversification through single-cell expression profiling

Antonio Scialdone^{1,2*}, Yosuke Tanaka^{3,4*†}, Wajid Jawaid^{3,4*}, Victoria Moignard^{3,4*}, Nicola K. Wilson^{3,4}, Iain C. Macaulay², John C. Marioni^{1,2,5} & Berthold Göttgens^{3,4}

In mammals, specification of the three major germ layers occurs during gastrulation, when cells ingressing through the primitive streak differentiate into the precursor cells of major organ systems. However, the molecular mechanisms underlying this process remain unclear, as numbers of gastrulating cells are very limited. In the mouse embryo at embryonic day 6.5, cells located at the junction between the extra-embryonic region and the epiblast on the posterior side of the embryo undergo an epithelial-to-mesenchymal transition and ingress through the primitive streak. Subsequently, cells migrate, either surrounding the prospective ectoderm contributing to the embryo proper, or into the extra-embryonic region to form the yolk sac, umbilical cord and placenta. Fate mapping has shown that mature tissues such as blood and heart originate from specific regions of the pre-gastrula epiblast¹, but the plasticity of cells within the embryo and the function of key cell-type-specific transcription factors remain unclear. Here we analyse 1,205 cells from the epiblast and nascent *Flk1*⁺ mesoderm of gastrulating mouse embryos using single-cell RNA sequencing, representing the first transcriptome-wide *in vivo* view of early mesoderm formation during mammalian gastrulation. Additionally, using knockout mice, we study the function of *Tal1*, a key haematopoietic transcription factor, and demonstrate, contrary to previous studies performed using retrospective assays^{2,3}, that *Tal1* knockout does not immediately bias precursor cells towards a cardiac fate.

Traditional experimental approaches for genome-scale analysis rely on large numbers of input cells and therefore cannot be applied to study early lineage diversification directly in the embryo. To address this, we used single-cell transcriptomics to investigate mesodermal lineage diversification towards the haematopoietic system in 1,205 single cells covering a time course from early gastrulation at embryonic day (E)6.5 to the generation of primitive red blood cells at E7.75 (Fig. 1a and Extended Data Figs 1a and 2a). Using previously published metrics (Methods), we observed that the data were of high quality. Five hundred and one single-cell transcriptomes were obtained from cells taken from dissected distal halves of E6.5 embryos sorted for viability only, which contain all of the epiblast cells, including the developing primitive streak, and a limited number of visceral endoderm and extra-embryonic ectoderm cells. From E7.0, embryos were staged according to anatomical features (Methods) as primitive streak, neural plate and head fold. The VEGF receptor *Flk1* (*Kdr*) was used to capture cells as it marks much of the developing mesoderm⁴. During subsequent blood development, *Flk1* is downregulated and *CD41* (*Itga2b*) is upregulated⁵. We therefore also sampled cells expressing both markers and *CD41* alone at the neural plate and head fold stages (Fig. 1a and Extended Data Figs 1b and 2a), giving a total of 138 cells from E7.0 (primitive streak), 259 from E7.5 (neural plate) and 307 from E7.75 (head fold).

After rigorous quality control, 2,085 genes were identified as having significantly more heterogeneous expression across the 1,205 cells than expected by chance (Extended Data Fig. 2b–d). Unsupervised hierarchical clustering in conjunction with a dynamic hybrid cut (Methods)

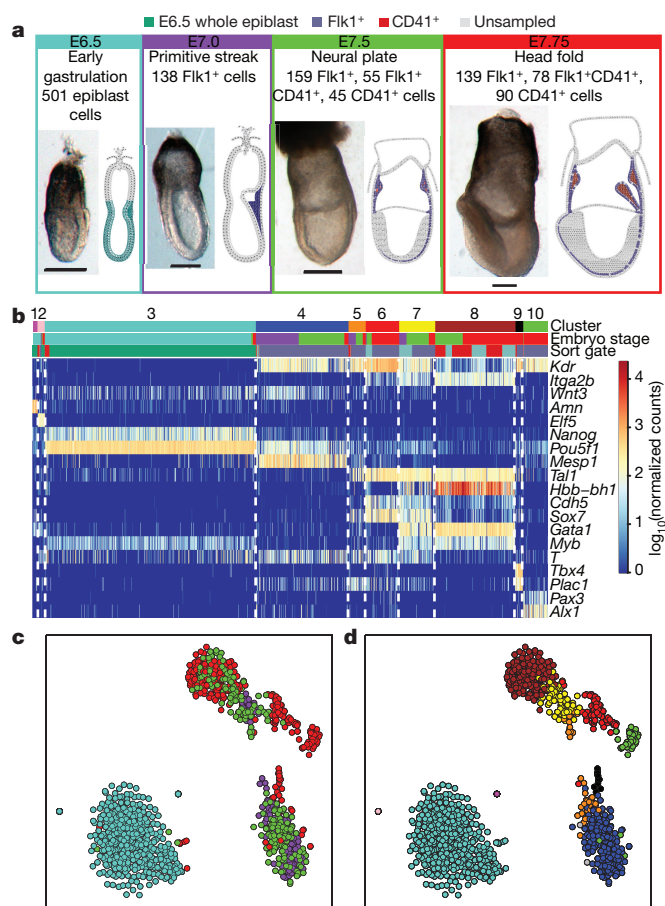


Figure 1 | Single-cell transcriptomics identifies ten populations relevant to early mesodermal development. **a**, Whole-mount images and schematics of E6.5–E7.75 embryo sections. Colours indicate approximate locations of sorted cells. Anterior, left; posterior, right. Scale bars, 200 μ m. **b**, Heatmap showing key genes distinguishing ten clusters. Coloured bars indicate assigned cluster (top), stage (middle: turquoise, E6.5; purple, primitive streak (E7.0); green, neural plate (E7.5); red, head fold (E7.75)) and the sorted population (bottom: green, E6.5 epiblast; blue, *Flk1*⁺; turquoise, *Flk1*⁺*CD41*⁺; red, *Flk1*⁺*CD41*⁺). **c**, t-SNE of all 1,205 cells coloured by embryonic stage, and **(d)** according to clusters in **b**.

¹EMBL-European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridge CB10 1SD, UK. ²Wellcome Trust Sanger Institute, Wellcome Genome Campus, Cambridge CB10 1SA, UK. ³Department of Haematology, Cambridge Institute for Medical Research, University of Cambridge, Cambridge CB2 0XY, UK. ⁴Wellcome Trust - Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK. ⁵Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge CB2 0RE, UK. [†]Present address: Division of Cellular Therapy, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan.

*These authors contributed equally to this work.

yielded ten robust clusters with varying contributions from the different embryonic stages (Fig. 1b, Extended Data Fig. 3, Methods and cell numbers in Extended Data Fig. 3h). Using t-distributed stochastic neighbour embedding (t-SNE) dimensionality reduction to visualize the data, three major groups were observed: one comprising almost all E6.5 cells, another mainly consisting of earlier primitive streak and neural plate stage cells, and a third containing predominantly later head fold stage cells (Fig. 1c). Importantly, clusters were coherent with the t-SNE visualization except for the small cluster 5 (Fig. 1d).

The expression of key marker genes allowed us to assign identities to each cluster: visceral endoderm, extra-embryonic ectoderm, epiblast, early mesodermal progenitors, posterior mesoderm, endothelium, blood progenitors, primitive erythrocytes, allantoic mesoderm and pharyngeal mesoderm (Fig. 1b, Extended Data Figs 3h and 4). Because of the limited cell numbers and lack of markers for their prospective isolation, conventional bulk transcriptome analysis of these key populations has never before been attempted.

Since the T-box transcription factor Brachyury—encoded by the *T* gene—marks the nascent primitive streak⁶, we investigated the gene expression programs associated with *T* induction in the E6.5 cells (cluster 3). *T* expression was restricted to a distinct subset of epiblast cells found closest to cluster 4 (Fig. 1d and Extended Data Fig. 5b), with rare isolated cells within the bulk of the epiblast population also expressing moderate levels, consistent with priming events for single gastrulation-associated genes. *T* expression correlated with other gastrulation-associated genes including *Mixl1* and *Mesp1* (Fig. 2a), with *Mesp1* highly expressed only in the small subset of cells situated at the pole of the E6.5 epiblast cluster (association of *T* and *Mesp1* expression: *P* value 3×10^{-15} , Fisher's exact test). We also observed a subset of cells distinct from the *T*⁺/*Mesp1*⁺ population, which expressed *Foxa2*, suggestive of endodermal priming⁷ (Extended Data Fig. 5d).

We next identified genes displaying correlated expression with *T*, which identified known markers and regulators such as *Mixl1*, and genes not previously implicated in mammalian gastrulation, such as *Slc35d3*, an orphan member of a nucleotide sugar transporter

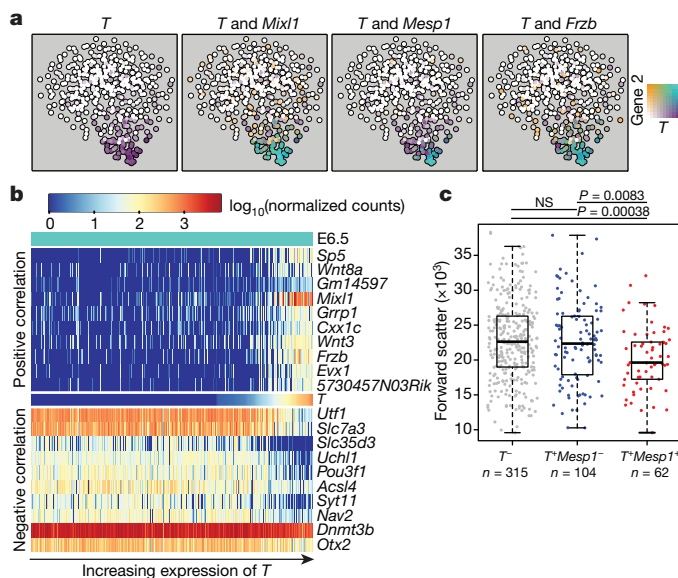


Figure 2 | Transcriptional program associated with *T* induction in E6.5 epiblast cells. **a**, t-SNE of the 481 E6.5 cells in cluster 3. Points are coloured by expression of *T* (Brachyury) and *Mixl1*, *Mesp1* and *Frzb*. **b**, Heatmap showing the ten genes most highly positively and negatively correlated with *T* (Supplementary Information Table 1). **c**, Forward scatter for the 481 E6.5 epiblast cells in cluster 3, with cells grouped according to *T*/*Mesp1* expression. Boxplots indicate the median and interquartile range. *P* values were calculated using a two-sided Welch's *t*-test for samples with unequal variance, with false discovery rate correction for multiple testing.

family⁸ and the retrotransposon-derived transcript *Cxx1c*⁹ (Fig. 2b and Supplementary Information Table 1). Genes negatively correlated with *T* were consistently expressed across the majority of epiblast cells, suggesting that cells outside the primitive streak have not yet committed to a particular fate, consistent with the known plasticity of epiblast cells in transplant experiments¹⁰. Ingressing epiblast cells undergo an EMT, turning from pseudo-stratified epithelial cells into individual motile cells, a conformational change associated with alterations in cell size and shape¹¹. Our E6.5 epiblast cells were isolated using index sorting, thus providing a forward scatter value for each cell. As shown in Fig. 2c, *T*⁺/*Mesp1*⁺ co-expressing cells showed a significant reduction in forward scatter values compared with *T*⁺/*Mesp1*⁻ and *T*⁻ cells. Since forward scatter correlates positively with cell size, this observation provides a direct link between specific transcriptional programs and characteristic physical changes associated with gastrulation. As *T*⁺/*Mesp1*⁺ cells also express *Mesp2*, this observation was consistent with the known EMT defect in *Mesp1*/*Mesp2* double knockout embryos¹². Index sorting therefore linked expression changes with dynamic physical changes similar to those recognized to occur during chicken gastrulation¹³.

We next focused on mesodermal lineage divergence during and immediately after gastrulation. We reasoned that approaches analogous to those used to order single cells in developmental pseudotime could be used to infer the location of cells in pseudospace, specifically with respect to the anterior–posterior axis of the primitive streak (Fig. 3a). To this end, we used diffusion maps¹⁴, a dimensionality reduction technique particularly suitable for reconstructing developmental trajectories¹⁵. We identified the diffusion-space direction that most probably represents true biological effects (see Methods), which we interpreted as the pseudospace coordinate (red line in Fig. 3b and Extended Data Fig. 6a–d). Hierarchical clustering revealed three groups of genes (Fig. 3c, Extended Data Fig. 6e and Supplementary Information Table 4) showing a gradient of expression along the pseudospace axis. These were assigned as anterior (darker blue, 334 genes) and posterior (lighter blue, 87 genes) owing to the enrichment of genes with known differential expression along the anterior–posterior axis of the primitive streak (Fig. 3d and Extended Data Figs 6f–h and 7). A third cluster was expressed highly at either end of the pseudospace axis (turquoise, 41 genes). Interestingly, the more posterior *Flkl1*⁺ mesodermal cells are associated with the allantois, blood and endothelial clusters (Fig. 1d and Extended Data Fig. 5c), which are known to arise from the posterior primitive streak. Gene ontology analysis revealed that the putative anterior genes were associated with terms relating to somite development, endoderm development and Notch signalling, consistent with a more anterior mesoderm identity¹⁶ (Supplementary Information Table 2a and Extended Data Fig. 6h). Conversely, the putative posterior mesoderm cluster was associated with BMP signalling, hindlimb development and endothelial cell differentiation, consistent with the posterior portion of the streak¹⁷.

Although derived from the same embryonic stages as the mesodermal progenitor cells, cluster 7 lacks expression of genes such as *Mesp1*, yet expresses *Tal1*, *Sox7*, *Tek* (Tie2) and *Fli1*, which are vital for extra-embryonic mesoderm formation (Fig. 1b and Extended Data Fig. 5, 7). Expression of *Kdr* and *Itga2b* (Extended Data Fig. 5b) further highlights clusters 7 and 8 (brown) as corresponding to the developmental journey towards blood, with a transition to mostly head fold stage cells in cluster 8 and increasing expression of embryonic haemoglobin *Hbb-bh1* (Fig. 1b). Given the apparent trajectory of blood development from cluster 7 to 8, we used an analogous approach to that described above to recover a pseudotemporal ordering of cells (Fig. 4a, Extended Data Fig. 8a–d and Methods). Eight hundred and three genes were downregulated, including the haematovascular transcription factor *Sox7*, which is known to be downregulated during blood commitment¹⁵ (Fig. 4c, d and Extended Data Fig. 8e, f). Sixty-seven genes were upregulated including the erythroid-specific transcription factors *Gata1* and *Nfe2*, and embryonic globin *Hbb-bh1*

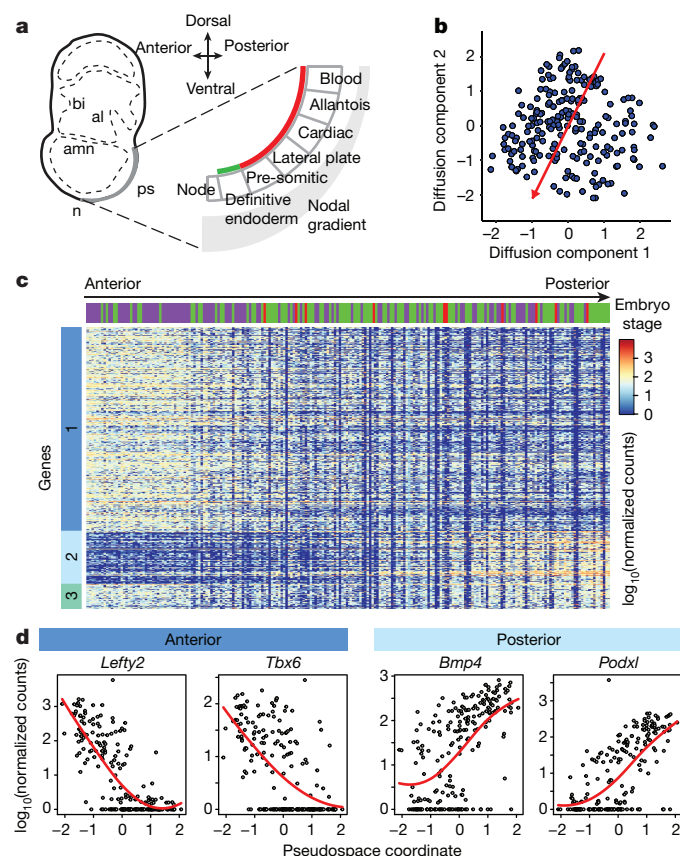


Figure 3 | Dimensionality reduction reveals transcriptional profiles associated with cell location in the embryo. **a**, Schematic of tissue emergence along the anterior–posterior primitive streak, derived from ref. 29. Mesodermally and endodermally derived tissues are marked by a red and green line, respectively; bi, blood island; al, allantois; amn, amnion; ps, primitive streak; n, node. **b**, Diffusion map of 216 cells in cluster 4 with pseudospace axis in red. Projections onto this axis represent pseudospace coordinates. **c**, Heatmap for differentially expressed genes along the pseudospace axis, showing genes more highly expressed in the anterior (dark blue) and posterior region (light blue), or highly expressed at either end (aquamarine). **d**, Expression profiles for example genes (red line, local polynomial fit).

(Fig. 4b, d, e and Extended Data Fig. 8). Twenty-seven genes were transiently expressed, including the known erythroid regulator *Gfi1b* (Supplementary Information Table 5). Significant GO terms associated with the upregulated genes were indicative of erythroid development, while downregulated genes were associated with other mesodermal processes including vasculogenesis and osteoblast differentiation (Supplementary Information Table 2b).

Gata1-null embryos die at around E10.5 owing to the arrest of yolk sac erythropoiesis¹⁸. We generated genome-wide ChIP-seq (chromatin immunoprecipitation followed by sequencing) data for Gata1 in haematopoietic cells derived after 5 days of embryonic stem (ESC) *in vitro* differentiation (Extended Data Fig. 9a–c). The group of upregulated genes from the pseudotime analysis showed a pronounced overlap with Gata1 targets ($P < 2.2 \times 10^{-16}$, Fisher's test) including known targets such as *Nfe2* and *Zfp1* (Fig. 4f, g, Extended Data Fig. 9d, e and Supplementary Information Table 6). Integration of single-cell transcriptomics with complementary transcription factor binding data therefore predicts likely *in vivo* targets of developmental regulators such as Gata1.

Two contrasting mechanisms are commonly invoked to explain how drivers of cell fate determination regulate cell type diversification. The first involves fate restriction through a stepwise sequence of binary fate choices and is supported by mechanistic investigations using ESC differentiation^{2,19}. The alternative invokes acquisition of diverse fates

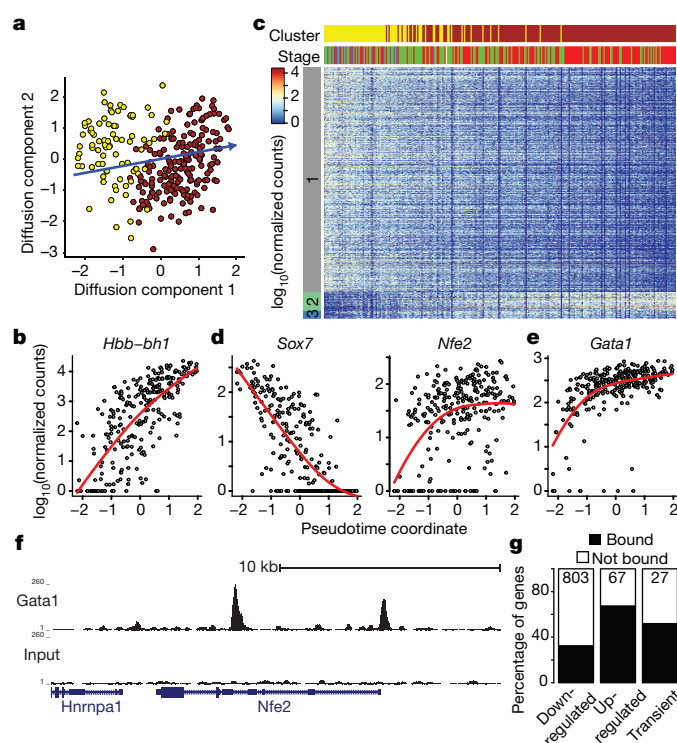


Figure 4 | Inferring the transcriptional program underlying primitive erythropoiesis. **a**, Diffusion map of 271 cells in clusters 7 and 8 displaying the inferred pseudotime axis (blue). **b**, Expression of *Hbb-bh1* ordered by pseudotime (red line, local polynomial fit). **c**, Heatmap ordered along the pseudotime axis. Horizontal bars indicate cluster and developmental stage. Genes shown were repressed (grey), activated (green) or transiently expressed (blue). **d**, Examples of activated and repressed genes and (e) *Gata1* as in **b**. **f**, University of California, Santa Cruz Browser tracks for Gata1 ChIP-seq and input in Runx1⁺ Gata1⁺ cells; the *Nfe2* locus is shown. **g**, Percentage of genes in each group identified in **c** overlapping Gata1 targets. Numbers indicate total numbers of genes in each category from **c**.

from independent precursor cells and is commonly supported by cell transplantation and lineage tracing analysis (Fig. 5a)^{1,10,20,21}. In contrast to the retrospective nature of transplantation and lineage tracing experiments where measurements are typically obtained a day or more after cell fate decisions are made, single-cell transcriptomics allows cellular states to be determined at the moment when fate decisions are executed since low cell numbers are not a limiting factor.

The bHLH transcription factor Tal1 (also known as Scl) is essential for the development of all blood cells^{22,23} with strong expression in posterior mesodermal derivatives (Fig. 5b). *Tal1*^{−/−} bipotential blood/endothelial progenitors cannot progress to a haemogenic endothelial state¹⁹, *Tal1* overexpression drives transdifferentiation of fibroblasts into blood progenitors²⁴ and *Tal1*^{−/−} mesodermal progenitors from the yolk sac give rise to aberrant cardiomyocyte progenitors when cultured *in vitro*². However, the precise nature of the molecular defect within *Tal1*^{−/−} mesodermal progenitors within the embryo has remained obscure, because cell numbers are too small for conventional analysis.

We profiled single Flk1⁺ cells from 4 wild type (WT) and 4 *Tal1*^{−/−} embryos obtained from E7.5 (neural plate) to E8.25 (four-somite stage) (256 WT and 121 *Tal1*^{−/−} cells; Fig. 5c and Extended Data Fig. 10), and computationally assigned cells to the previously defined 10 clusters (Methods). Cells from WT embryos contributed to all clusters, while *Tal1*^{−/−} embryos did not contain any cells corresponding to the blood progenitor and primitive erythroid clusters (yellow and brown, Fig. 5d) consistent with the known failure of primitive erythropoiesis in *Tal1*^{−/−} embryos²³ and their lack of CD41 expression (Fig. 5c).

Forty-five *Tal1*^{−/−} cells were confidently mapped to the endothelial (red) cluster, which therefore allowed us to investigate the early consequences of *Tal1* deletion in this key population for definitive

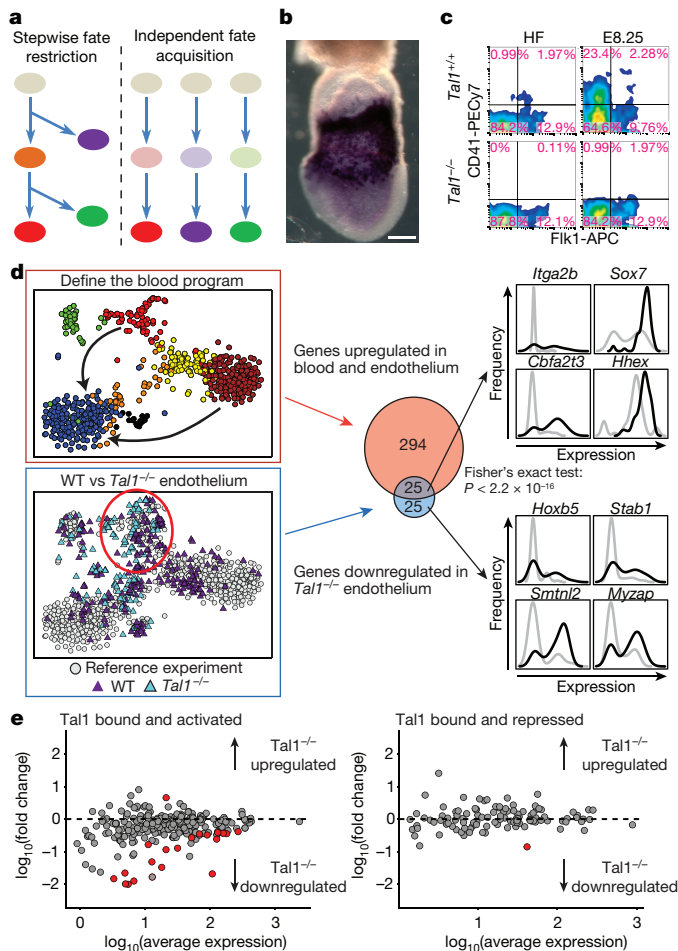


Figure 5 | Analysis of *Tal1*^{-/-} embryos suggests independent fate acquisition. **a**, Two cell fate diversification models. **b**, *Tal1* in situ hybridization at head fold stage. Scale bar, 200 μ m. **c**, Flow cytometry of WT and *Tal1*^{-/-} mice at head fold and E8.25. **d**, Blood program genes are differentially expressed between nascent mesoderm (blue) and endothelial (red) and blood cells (brown). Differential expression between 45 *Tal1*^{-/-} and 59 WT endothelial cells (lower left t-SNE) identified 50 downregulated genes. Gene set overlap (centre) indicates failure to induce the blood program in *Tal1*^{-/-} endothelium ($P < 2.2 \times 10^{-16}$, Fisher's test). On the right are expression distributions for selected genes in WT (black) or *Tal1*^{-/-} (grey) endothelial cells. **e**, For genes previously reported³ to be bound and activated (left) or bound and repressed (right) by *Tal1*, fold change between *Tal1*^{-/-} and WT endothelium (defined in **d**) is plotted against average expression. Red circles, genes with a fold change > 1.5 and a false discovery rate < 0.05 .

haematopoietic development (Fig. 5d and Supplementary Information Tables 7 and 8). Fifty genes were downregulated in *Tal1*^{-/-} endothelial cells (fold change < 0.67 , 5% false discovery rate). These included known regulators of early blood development (*Itga2b*, *Lyl1*, *Cbfa2t3*, *Hhex*, *Fli1*, *Ets2*, *Egfl7*, *Sox7*, *Hoxb5*), consistent with *Tal1* specifying a haematopoietic fate in embryonic endothelial progenitor cells¹⁹, and in particular *Hoxb5*, which has recently emerged as a powerful marker for definitive blood stem cells²⁵. Single-cell profiling also identified genes with altered distributions of expression. For example, *Sox7* changed from a largely unimodal pattern in WT cells to a bimodal on/off pattern in *Tal1*^{-/-} endothelial cells, while *Cbfa2t3* showed the opposite pattern (Fig. 5d).

However, we did not observe upregulation of cardiac markers in *Tal1*^{-/-} endothelial cells (Fig. 5e and Supplementary Information Tables 8 and 9). Previously, this upregulation had been observed in yolk sac endothelial cells collected 1–1.5 days later than our data², and had been taken as evidence that *Tal1* acts as a gatekeeper controlling

the balance between alternative cardiac and blood/endothelial fates within single multipotent mesodermal progenitors³. Our results, however, suggest that the primary role of *Tal1* is induction of a blood program, and the subsequent ectopic expression of cardiac genes may be the result of secondary induction events acting on a still relatively plastic mesodermal cell blocked from executing its natural developmental program.

Here we have used single-cell transcriptomics to obtain a comprehensive view of the transcriptional programs associated with mammalian gastrulation and early mesodermal lineage diversification. Further technological advances to resolve epigenetic processes at single-cell resolution²⁶ and match single-cell expression profiles with spatial resolution^{27,28} are probably key drivers of future progress in this field. Finally, our analysis of *Tal1*^{-/-} embryos illustrates how the phenotypes of key regulators can be re-evaluated at single-cell resolution to advance our understanding of early mammalian development.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 March; accepted 9 June 2016.

Published online 6 July 2016.

- Lawson, K. A., Meneses, J. J. & Pedersen, R. A. Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* **113**, 891–911 (1991).
- Van Handel, B. *et al.* Scl represses cardiomyogenesis in prospective hemogenic endothelium and endocardium. *Cell* **150**, 590–605 (2012).
- Org, T. *et al.* Scl binds to primed enhancers in mesoderm to regulate hematopoietic and cardiac fate divergence. *EMBO J.* **34**, 759–777 (2015).
- Ema, M. *et al.* Primitive erythropoiesis from mesodermal precursors expressing VE-cadherin, PECAM-1, Tie2, endoglin, and CD34 in the mouse embryo. *Blood* **108**, 4018–4024 (2006).
- Mikkola, H. K. A., Fujiwara, Y., Schlaeger, T. M., Traver, D. & Orkin, S. H. Expression of CD41 marks the initiation of definitive hematopoiesis in the mouse embryo. *Blood* **101**, 508–516 (2003).
- Wilkinson, D. G., Bhatt, S. & Herrmann, B. G. Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* **343**, 657–659 (1990).
- Burtscher, I. & Lickert, H. Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development* **136**, 1029–1038 (2009).
- Chintala, S. *et al.* The Slc35d3 gene, encoding an orphan nucleotide sugar transporter, regulates platelet-dense granules. *Blood* **109**, 1533–1540 (2007).
- Henke, C. *et al.* Selective expression of sense and antisense transcripts of the sushi-ichi-related retrotransposon – derived family during mouse placental development. *Retrovirology* **12**, 9 (2015).
- Tam, P. P. L. & Zhou, S. X. The allocation of epiblast cells to ectodermal and germ-line lineages is influenced by the position of the cells in the gastrulating mouse embryo. *Dev. Biol.* **178**, 124–132 (1996).
- Solnica-Krezel, L. & Sepich, D. S. Gastrulation: making and shaping germ layers. *Annu. Rev. Cell Dev. Biol.* **28**, 687–717 (2012).
- Kitajima, S., Takagi, A., Inoue, T. & Saga, Y. MesP1 and MesP2 are essential for the development of cardiac mesoderm. *Development* **127**, 3215–3226 (2000).
- Rozbicki, E. *et al.* Myosin-II-mediated cell shape changes and cell intercalation contribute to primitive streak formation. *Nature Cell Biol.* **17**, 397–408 (2015).
- Haghverdi, L., Büttner, F. & Theis, F. J. Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics* **31**, 2989–2998 (2015).
- Moignard, V. *et al.* Decoding the regulatory network of early blood development from single-cell gene expression measurements. *Nature Biotechnol.* **33**, 269–276 (2015).
- Saga, Y. Segmental border is defined by the key transcription factor *Mesp2*, by means of the suppression of Notch activity. *Dev. Dyn.* **236**, 1450–1455 (2007).
- Lawson, K. A. *et al.* Brmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev.* **13**, 424–436 (1999).
- Fujiwara, Y., Browne, C. P., Cunneiff, K., Goff, S. C. & Orkin, S. H. Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc. Natl Acad. Sci. USA* **93**, 12355–12358 (1996).
- Lancrin, C. *et al.* The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. *Nature* **457**, 892–895 (2009).
- Padrón-Barthe, L. *et al.* Clonal analysis identifies hemogenic endothelium and not hemangioblasts as the source of the blood-endothelial common lineage in the mouse embryo. *Blood* **124**, 2523–2532 (2014).
- Tam, P. P., Parameswaran, M., Kinder, S. J. & Weinberger, R. P. The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. *Development* **124**, 1631–1642 (1997).
- Porcher, C. *et al.* The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell* **86**, 47–57 (1996).

23. Shivdasani, R. A., Mayer, E. L. & Orkin, S. H. Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature* **373**, 432–434 (1995).
24. Batta, K., Florkowska, M., Kouskoff, V. & Lacaud, G. Direct reprogramming of murine fibroblasts to hematopoietic progenitor cells. *Cell Reports* **9**, 1871–1884 (2014).
25. Chen, J. Y. *et al.* Hoxb5 marks long-term haematopoietic stem cells and reveals a homogenous perivascular niche. *Nature* **530**, 223–227 (2016).
26. Bheda, P. & Schneider, R. Epigenetics reloaded: the single-cell revolution. *Trends Cell Biol.* **24**, 712–723 (2014).
27. Achim, K. *et al.* High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nature Biotechnol.* **33**, 503–509 (2015).
28. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnol.* **33**, 495–502 (2015).
29. Robertson, E. J. Dose-dependent Nodal/Smad signals pattern the early mouse embryo. *Semin. Cell Dev. Biol.* **32**, 73–79 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. de Bruijn, A. Martinez-Arias, J. Nichols and C. Mulas for discussion, the Cambridge Institute for Medical Research Flow Cytometry facility for their expertise in single-cell index sorting, and S. Lorenz from the Sanger Single Cell Genomics Core for supervising purification of *Tal1*^{-/-} sequencing libraries. ChIP-seq reads were processed by R. Hannah. Research in the authors' laboratories is supported by the Medical Research Council, Cancer Research UK, the Biotechnology and Biological Sciences Research Council, Bloodwise, the Leukemia and Lymphoma Society, and the Sanger-EBI Single Cell Centre, and by core support grants from the Wellcome Trust to the Cambridge Institute for Medical Research and Wellcome

Trust - MRC Cambridge Stem Cell Institute and by core funding from Cancer Research UK and the European Molecular Biology Laboratory. Y.T. was supported by a fellowship from the Japan Society for the Promotion of Science. W.J. is a Wellcome Trust Clinical Research Fellow. A.S. is supported by the Sanger-EBI Single Cell Centre. This work was funded as part of Wellcome Trust Strategic Award 105031/D/14/Z 'Tracing early mammalian lineage decisions by single-cell genomics' awarded to W. Reik, S. Teichmann, J. Nichols, B. Simons, T. Voet, S. Srinivas, L. Vallier, B. Göttgens and J. Marioni.

Author Contributions A.S. and W.J. processed and analysed single-cell RNA sequencing (RNA-seq) data. A.S. and V.M. generated figures. Y.T. and W.J. performed embryo dissection. N.K.W., V.M. and J.C.M. performed single-cell RNA-seq experiments. Y.T. performed flow cytometry, ESC differentiation and *in situ* hybridization. V.M. performed ChIP-seq assays. A.S., W.J., Y.T., V.M., B.G. and J.C.M. interpreted results and wrote the paper. B.G. and J.C.M. supervised and conceived the study.

Author Information ChIP-seq data are available at the NCBI Gene Expression Omnibus portal under accession number GSE74994. Processed data are also available at <http://codex.stemcells.cam.ac.uk>. RNA-seq data are available at Array Express under accession numbers E-MTAB-4079 and E-MTAB-4026. Processed RNA-seq data are also available at <http://gastrulation.stemcells.cam.ac.uk/scialdone2016>. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.G. (bg200@cam.ac.uk) or J.C.M. (marioni@ebi.ac.uk).

Reviewer Information *Nature* thanks A.-K. Hadjantonakis, P. Robson and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Timed matings and embryo collection. All procedures were performed in strict adherence to United Kingdom Home Office regulations (project licence 70/8406). Timed matings were set up between CD1 mice (which produce large litters). Embryos were staged according to the morphological criteria of Downs and Davies³⁰, and classified broadly as primitive streak, neural plate or head fold stage. Suspensions of cells from individual embryos were prepared by incubating with TrypLE Express dissociation reagent (Life Technologies) at 37°C for 10 min and quenching with heat-inactivated serum. All cells were stained with DAPI for viability. At E6.5, the distal half of the embryo was dissected and dissociated into a single-cell suspension, and live cells were sorted. For E7.0 and older, suspensions consisted of the whole embryo and were also stained with Flk1-APC (AVAS12 at 1:400 dilution; BD Bioscience) and only Flk1⁺ cells were collected. For cell sorting of CD41⁺Flk1⁺ and CD41⁺Flk1⁺ cells from neural plate stage and head fold stages, suspensions were stained with Flk1-APC, PDGFRa-PE (APA5 at 1:200 dilution; Biolegend) and CD41-PEcy7 (MWReg30 at 1:400 dilution; Biolegend) for 20 min at 4°C as described³¹. Cells were sorted from seven E6.5 embryos. Flk1⁺ cells were sorted from three primitive streak stage, four neural plate stage and three head fold stage embryos (Extended Data Fig. 1a). CD41⁺Flk1⁺ and CD41⁺Flk1⁺ cells were sorted from the same embryos, an additional eight each at neural plate and head fold stages (Extended Data Fig. 1b). Cell sorting was performed with a BD Influx cell sorter in single-cell sort mode with index sorting to confirm the presence of a single event in each well. Additional cells were sorted into tissue culture plates to visually confirm the presence of single events.

To obtain *Tal1*^{-/-} cells, timed matings were set up between *Tal1*^{LacZ/+} mice³². Flk1⁺ cells were sorted as above from four embryos for each genotype: from one embryo for each genotype at neural plate and four-somite (4S) stages, from two head fold stage embryos for *Tal1*^{LacZ/LacZ} (designated *Tal1*^{-/-}), one head fold stage WT embryo and one WT embryo intermediate between neural plate and head fold stages. Genotyping PCR using 1/20 suspension cells was performed as described previously³².

Single-cell RNA sequencing library preparation and mapping of reads. scRNA-seq analysis used the Smart-seq2 protocol as previously described³³. Single cells were sorted by fluorescence-activated cell sorting (FACS) into individual wells of a 96-well plate containing lysis buffer (0.2% (v/v) Triton X-100 and 2 U/μl RNase inhibitor (Clontech)) and stored at -80°C. Libraries were prepared using the Illumina Nextera XT DNA preparation kit and pooled libraries of 96 cells were sequenced on the Illumina Hi-Seq 2500. Reads were mapped simultaneously to the *Mus musculus* genome (Ensembl version 38.77) and the ERCC sequences using GSNAP (version 2014-10-07) with default parameters. HTseq-count³⁴ was used to count the number of reads mapped to each gene (default options).

Identification of poor quality cells. To assess data quality³⁵, five metrics were used: (1) total number of mapped reads, (2) fraction of total reads mapped to endogenous genes, (3) fraction of reads mapped to endogenous genes that are allocated to mitochondrial genes, (4) fraction of total reads mapped to ERCC spike-ins and (5) level of sequence duplication (as estimated by FastQC, version 0.11.4, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>).

For all downstream analyses, we only retained samples that had (1) more than 200,000 reads mapped (either to ERCC spike-ins or endogenous mRNA), (2) more than 20% of total reads mapped to mRNA, (3) less than 20% of mapped reads allocated to mitochondrial genes, (4) less than 20% of reads mapped to ERCC spike-ins and (5) less than 80% of duplicated sequences. Out of the 2,208 cells that were captured across the two experiments, 1,582 (that is, ~72% of the total) passed our quality check. A t-SNE projection³⁶ of the values of these five metrics (Extended Data Fig. 2b) shows that most of discarded cells tend to cluster together and fail at least two criteria. All metrics were standardized before applying t-SNE with the 'RtSNE' function (default parameters) from the R package 'RtSNE' (version 0.1)³⁷.

Normalization of read counts. The data were normalized for sequencing depth using size factors³⁸ calculated on endogenous genes. By doing so, we also normalized for the amount of RNA obtained from each cell³⁹, which is itself highly correlated with cell cycle stage⁴⁰.

Highly variable genes and GO enrichment analysis. Highly variable genes were identified by using the method described in Brenneke *et al.*³⁹. In brief, we fitted the squared coefficient of variation as function of the mean normalized counts³⁹. In the fitting procedure, to minimize the skewing effect due to the lowly expressed genes³⁹, only genes with a mean normalized count greater than 10 were used. Genes with an adjusted *P* value (Benjamini-Hochberg method) less than 0.1 were considered significant (red circles in Extended Data Fig. 2c). This set of highly variable genes was used for the clustering analysis discussed below. The GO enrichment analysis was performed using TopGO in its 'elimination mode' with Fisher's

exact test; we considered as significant GO categories with an unadjusted *P* value below 10⁻⁴.

Differentially expressed genes. To find genes differentially expressed between two groups of cells we used edgeR⁴¹ (version 3.12). Before running edgeR, we excluded genes annotated as pseudogenes in Ensembl, sex-related genes (Xist and genes on the Y chromosome) and genes that were not detected or were expressed at very low levels (we considered only genes that had more than ten reads per million in at least *n* cells, *n* being equal to 10% of the cells in the smaller group being compared). The function 'glmTreat' was then used to identify the genes having a fold change significantly greater than 1.5 at a false discovery rate threshold equal to 0.05.

Clustering analysis. Clustering analysis was performed on the 1,205 WT cells from the first experiment that passed the QC. The Spearman correlation coefficient, ρ , was computed between the expression levels of highly variable genes in each pair of cells, which was then used to build a dissimilarity matrix defined as $(1 - \rho)/2$. Hierarchical clustering was performed ('hclust' R function with the 'average' method) on the dissimilarity matrix and clusters were identified by means of the dynamic hybrid cut algorithm⁴². The R function 'cutreeDynamic' with the 'hybrid' method and a minimum cluster size equal to ten cells was used ('dynamicTreeCut' package, version 1.62). This function allows the user to specify the 'deepSplit' parameter that controls the sensitivity of the method: higher values of this parameter correspond to higher sensitivity and can result in more clusters being identified, but also entail an increased risk of overfitting the data. The optimal trade-off between robustness of clustering and sensitivity was found by analysing the results of the algorithm with all possible values of the deepSplit parameter (that is, integer values from 0 to 4) on 100 subsamples of our data. In particular, in each subsample, we removed 10% of genes randomly selected before computing the dissimilarity matrix and applying the clustering algorithm.

The statistics of the Pearson gamma and the average silhouette width (computed with the 'cluster.stats' function included in the R package 'fpc', version 2.1-10)^{43,44} of the subsamples (see Extended Data Fig. 3a,b) suggest that with 'deepSplit=2' a good compromise is reached between robustness and sensitivity for our data. We identified ten different clusters as well as two outlier cells that, although similar in gene expression to the mesodermal progenitor cells (cluster 4), were not assigned to any cluster by the algorithm, probably because of their relatively poor quality.

We then evaluated how specifically each gene is expressed in any given cluster. First, we found the differentially expressed genes (as described above) between all pairs of clusters. Marker genes for cluster *i* are expected to be significantly upregulated in *i* across all pairwise comparisons involving cluster *i*. The average rank of a marker gene across the pairwise comparisons provides a measure of how specifically the marker is expressed in the cluster. Extended Data Fig. 3c-f shows the expression values of marker genes for four different clusters. We provide the full list of markers in Supplementary Information Table 3. The clusters were visualized by using t-SNE (as implemented in the 'RtSNE' R package) on the dissimilarity matrix. **Single-cell trajectories in pseudospace: the anterior/posterior axis of the primitive streak.** As discussed in the main text, cells allocated to cluster 4 (Fig. 1b-d) are cells that have probably exited the primitive streak only recently. We sought to align the cells along a pseudospacial trajectory representing the anterior-posterior axis of the primitive streak, which would allow us to identify the likely original locations of each cell along such an axis.

To do this we adopted an unsupervised approach: we did not use any prior information about marker genes, but selected the strongest signal present in this cluster of cells (controlling for potential batch effects) and later verified its biological meaning. We first used a diffusion map-based technique to reduce the dimensionality of the data set. Diffusion maps have recently been successfully applied to identify developmental trajectories in single-cell qPCR and RNA-seq data^{14,15}. We used the implementation of the 'destiny' R package ('DiffusionMap' function) developed by Angerer *et al.*⁴⁵. We restricted the analysis to genes that are highly variable among cells in the blue cluster and have an average expression above ten normalized read counts. The centred cosine similarity was used ('cosine' option in the 'DiffusionMap' function) and only the first two diffusion components (DC1 and DC2) were retained for downstream analysis.

In addition to biologically meaningful signals, batch effects (owing to cells being sorted and processed on different plates) can also be present and induce structure within the data. While in our data set the batch effect does not strongly influence the definition of different populations of cells, it might become relevant when finer structures within a single cluster of cells are considered (see Extended Data Fig. 6a). To tease apart the signals due to biological and batch effects, we computed the fraction of variance attributable to the batch effect along each direction in the diffusion space using a linear regression model. The direction 'orthogonal' to the batch effect, that is, the direction associated with the smallest fraction of variance explained by the batch effect, was considered as mostly driven by a biologically relevant signal. Hence, all cells were projected on this direction to obtain a 'pseudo-coordinate' representing the state of a cell relative to the biological process

captured by the diffusion map. The direction was identified by the angle α that it formed with the DC1 axis (Extended Data Fig. 6c).

Cells considered here are mostly from two batches including cells from the primitive streak stage (plate SLX-8408 and SLX-8409) and two batches including cells from the neural plate stage (plate SLX-8410 and SLX-8411; Extended Data Fig. 6b). For each of these two sets of batches, we computed the fraction of variance that can be explained by the batch covariate along any possible direction in the diffusion plot by using a linear regression model. The angles α_1 and α_2 corresponding to the directions orthogonal to the two batch effects are very close to each other (Extended Data Fig. 6c); we took the average value of α between these two angles to approximate the direction orthogonal to both batch effects.

Cells' coordinates in the diffusion space were projected along the direction identified by the average value of α , and this projection was interpreted as a 'pseudospace' coordinate representing the position of cells along the primitive streak (see main text and Fig. 3). We tested the robustness of such a pseudospace coordinate by repeating the same analysis with alternative dimensionality reduction techniques (t-SNE and independent component analysis), which gave highly correlated coordinates (see Extended Data Fig. 6d). A principal component analysis performed with a set of previously known markers for the anterior and posterior regions of the primitive streak also yielded a first component highly correlated with the pseudospace coordinate (see Extended Data Fig. 6h left panel). Moreover, the pseudospace coordinate had a positive (negative) correlation with the posterior (anterior) markers used (see Extended Data Fig. 6h right panel). These results strongly support the robustness of the signal we identified as well as its biological interpretation.

Once the pseudospace trajectory was defined, we selected genes that were differentially expressed along the trajectory. First, we removed all genes that were not detected in any cell. Then, for each gene, we fitted the $\log_{10}(\text{expression levels})$ (adding a pseudocount of 1) by using two local polynomial models: one with degree 0 and another with degree 2 ('loclfit' function in 'loclfit' R package, nearest neighbour component parameter equal to 1). The first, simpler model is better suited for genes that do not change their expression level along the trajectory. The second model has a greater number of parameters and is able to reproduce the more complex dynamics of genes that are differentially expressed.

We evaluated these two models by using the Akaike information criterion (AIC), a score that measures how well the data are reproduced by the model and includes a penalization for more complex models⁴⁶. Better models according to this criterion correspond to smaller AIC scores.

To compute the AIC scores for the two models, we used the 'aic' function available in the 'loclfit' R package, and then calculated the difference: $\Delta\text{AIC} = \text{AIC}(\text{degree} = 2) - \text{AIC}(\text{degree} = 0)$. Negative values indicate that the more complex model with degree 2 local polynomials performs better, and therefore corresponds to genes that are more likely to be differentially expressed. Genes having a $\Delta\text{AIC} < -2$ were considered to be significantly differentially expressed along the trajectory⁴⁶.

A hierarchical tree was built with the normalized expression patterns of the 462 differentially expressed genes (function 'hclust' with average linkage method and dissimilarity based on Spearman correlation) and a dynamic hybrid cut algorithm ('cutreeDynamic' function, minimum cluster size equal to 5) split this set of genes into three clusters according to the type of dynamics they have (see Fig. 3, Extended Data Fig. 6e and Supplementary Information Table 4).

Single-cell trajectories in pseudotime: the blood developmental trajectory. As discussed in the main text, clusters 7 and 8 (yellow and brown clusters in Fig. 1b, d) include blood progenitors at different stages of differentiation. By using a procedure analogous to the one described above, we aligned these cells along a trajectory representing embryonic blood development.

Extended Data Fig. 8a shows the diffusion plot with cells from the yellow and the brown clusters. Most of these cells come from plates SLX-8344 and SLX-8345 that were collected from embryos at neural plate and late head fold stages (see Extended Data Fig. 8b). With a linear regression model, where we controlled for biological parameters such as stage and sorting, we found the direction that correlates the least with the batch effect associated to these two plates and projected all cells onto it (Extended Data Fig. 8c). Note that the minimum correlation with the batch effect is achieved at a very small value of α ($\sim 10^\circ$, see Extended Data Fig. 8c), suggesting that the first diffusion component is mainly driven by a biologically meaningful signal and the batch effect plays a minor role here even at this more detailed scale of analysis. The new cell coordinate obtained from the projection was interpreted as a 'pseudotime' coordinate, which represents the differentiation stage of each cell along their journey towards erythroid fate. As expected, cells in the yellow cluster have a smaller pseudotime coordinate compared with the brown cluster that is mainly composed of more differentiated primitive erythroid cells. An analysis with alternative dimensionality reduction techniques yielded highly correlated pseudotime coordinates, suggesting the robustness of the signal (Extended Data Fig. 8d). Furthermore, our biological interpretation of the pseudotime coordinate

is supported by the expression pattern of genes that are known to be upregulated or downregulated along the blood developmental trajectory, as is clear via principal component analysis (see Extended Data Fig. 8f).

By using the filtering and clustering procedure described in the previous section, we were able to detect 897 genes that were differentially expressed along the trajectory, which were divided in three clusters, each displaying a different type of dynamics (see Extended Data Fig. 8e and Supplementary Information Table 5). **Random Forest to allocate cells to previously identified clusters.** Cells captured in the Tal1 experiment (testing data set) were allocated to the clusters we previously identified by using a Random Forest algorithm⁴⁷ (R package 'randomForest', version 4.6-12)⁴⁸ trained on the cells captured in the first experiment (training data set). The rank-normalized expression levels of all highly variable genes in the training data set were used as variables (the R function 'rank' was used for normalization, ties were averaged). The Random Forest algorithm was first used on the training data to assess variable importance with 1,000 classification trees. The 25% most important variables were selected to grow another set of 1,000 trees that were then used for the classification of the testing data set. With this filtered set of variables, the out-of-bag error estimate was $\sim 4.8\%$.

The quality of allocation of each cell in the testing data set was verified by computing the median of pairwise dissimilarities (defined as $(1 - \rho)/2$, with ρ being the Spearman correlation) of that cell to all other cells in the training data allocated to the same cluster. Cells in the testing data set having a median pairwise dissimilarity larger than the maximum of the medians of pairwise dissimilarities of cells in the training data were considered to be 'unclassified' ($\sim 1.8\%$ of all cells from the testing data set). For the identification of differentially expressed genes between clusters in the testing data, only cells that were confidently allocated to the clusters (that is, cells with a minimum difference of 10% probability between the best and the second best cluster allocation) were used.

Generation, maintenance and haematopoietic differentiation of Runx1-GFP/Gata1-mCherry ESCs. *Runx1^{GFP/+}Gata1^{mCherry/+}* ESCs were generated from morulae as described previously^{49,50}. Cells were not tested for mycoplasma contamination. ESCs were grown on gelatinized plates (0.1% gelatin in water) at 37°C and 5% CO₂ in ESC media (Knockout DMEM (Life Technologies) with 15% FCS (batch-tested for ESC culture; Life Technologies), 2 mM L-glutamine (PAA Laboratories), 0.5% P/S, 0.1 mM β -mercaptoethanol (Life Technologies) and 10³ U/ml recombinant LIF (ORF Genetics)). Cells were passaged with TrypLE Express dissociation reagent (Life Technologies) every 1–3 days. ESCs were differentiated as embryoid bodies as previously described^{31,51}. Embryoid bodies were harvested into Falcon tubes after 5 days of culture and dissociated with TrypLE Express dissociation reagent and prepared for FACS.

ChIP-seq. ChIP was performed as described⁵² with modifications for low cell numbers⁵³. Approximately 7×10^6 FACS-sorted day 5 embryoid body cells (*Runx1-ires-GFP⁺/Gata1-mCherry⁺*; Extended Data Fig. 9a) per ChIP were cross-linked using formaldehyde to a final concentration of 1%. As samples were pooled from several sorts, isolated nuclei were frozen on dry ice-cold isopropanol and stored at -80°C . During the immunoprecipitation step, 4 μl recombinant histone 2B (New England Biolabs) and 1 μl of mouse RNA (Qiagen; diluted 1/5 in IP dilution buffer) were added as carriers, followed by 7 μg of primary antibody (rabbit anti-Gata1, Abcam ab11963). Sequencing libraries were prepared using the TruSeq Kit (Illumina) for high throughput sequencing on an Illumina HiSeq 2500, according to the manufacturer's instructions, with size selection for fragments of 150–400 bp. **ChIP-seq mapping and analysis.** Alignment of the ChIP-seq reads to the mouse mm10 genome, quality control and peak calling were performed according to the data pipeline set out by Sanchez-Castillo *et al.*⁵⁴. Peak calling was performed using MACS2⁵⁵ with $P = 1 \times 10^{-6}$. Post-processing using in-house scripts converted the peak coordinates to 400 bp on the basis of peak summits given in the MACS output. Coordinates of genomic regions that lie at the end of chromosomes and/or in repeat regions were discarded from the final high-confidence peak lists. PolyApeak⁵⁶ was run in R to remove abnormally shaped peaks. Peaks were assigned to genes using an in-house script according to whether they overlapped with a known TSS or fell within 50 kb each side of a gene.

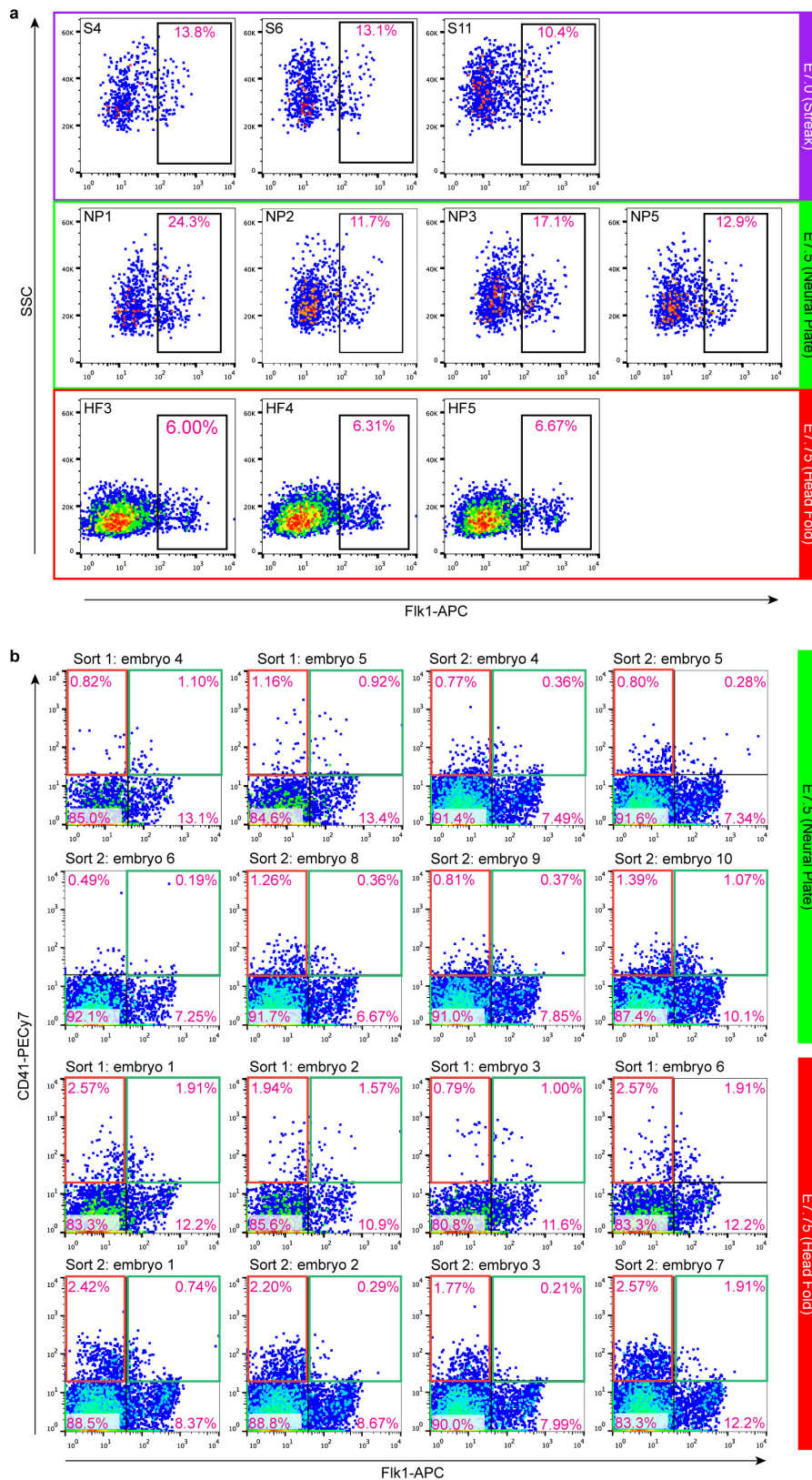
In situ hybridization. Whole-mount *in situ* hybridization for *Tal1* was performed as described previously⁵⁷. An *in situ* hybridization probe for *Tal1* was synthesized using published sequence (*Tal1* 860–1428, accession number M59764) with the DIG RNA labelling kit (Roche).

Code availability. All data were analysed with standard programs and packages, as detailed above. Code is available on request.

30. Downs, K. M. & Davies, T. Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope. *Development* **118**, 1255–1266 (1993).
31. Wilkinson, A. C. *et al.* Single site-specific integration targeting coupled with embryonic stem cell differentiation provides a high-throughput alternative to *in vivo* enhancer analyses. *Biol. Open* **2**, 1229–1238 (2013).

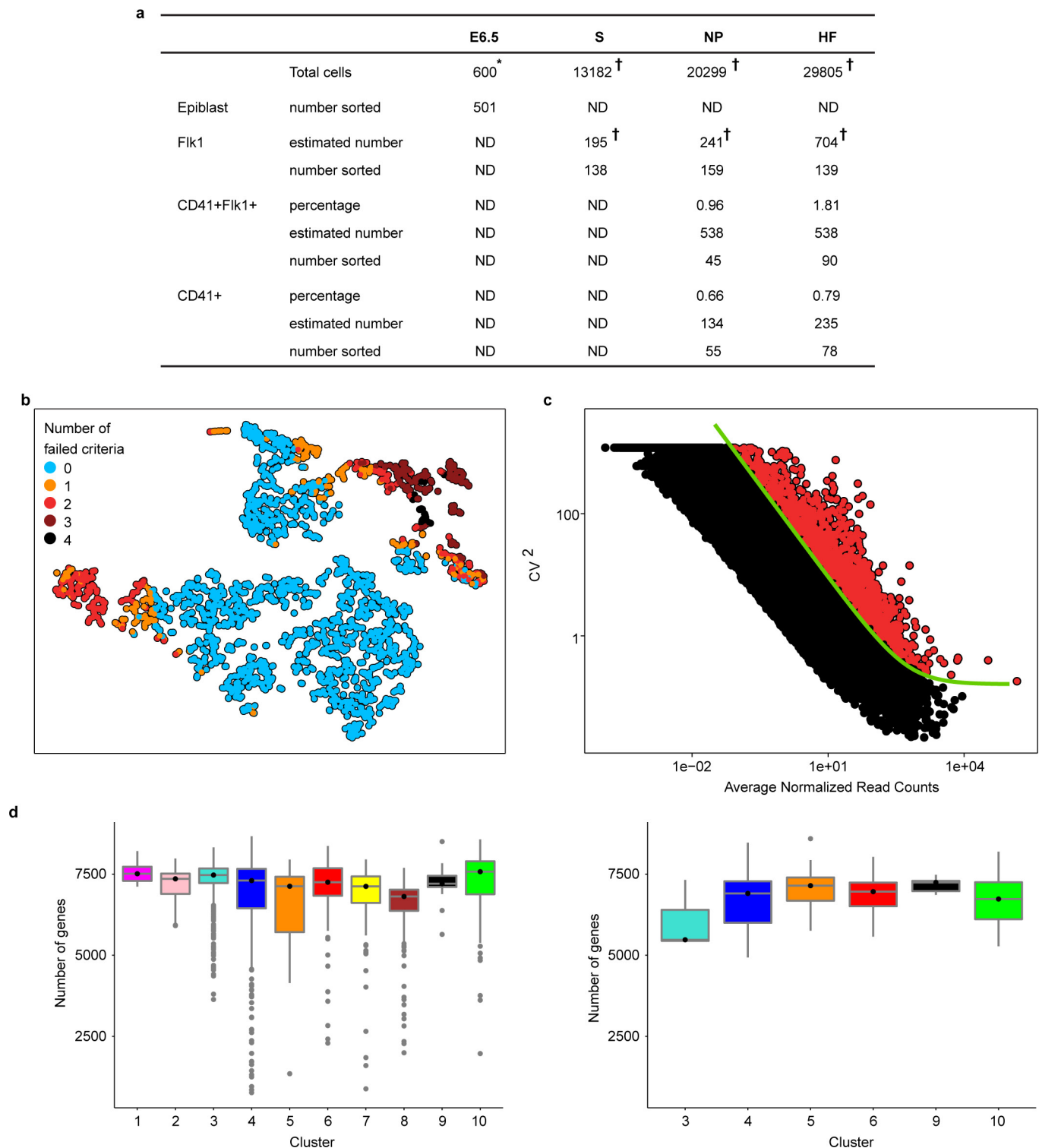
32. Elefanti, A. G. *et al.* Characterization of hematopoietic progenitor cells that express the transcription factor SCL, using a lacZ “knock-in” strategy. *Proc. Natl Acad. Sci. USA* **95**, 11897–11902 (1998).
33. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols* **9**, 171–181 (2014).
34. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
35. Stegle, O., Teichmann, S. A. & Marioni, J. C. Computational and analytical challenges in single-cell transcriptomics. *Nature Rev. Genet.* **16**, 133–145 (2015).
36. van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
37. van der Maaten, L. Barnes-Hut-SNE. Preprint at <http://arxiv.org/pdf/1301.3342> (2013).
38. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
39. Brennecke, P. *et al.* Accounting for technical noise in single-cell RNA-seq experiments. *Nature Methods* **10**, 1093–1095 (2013).
40. Buettner, F. *et al.* Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nature Biotechnol.* **33**, 155–160 (2015).
41. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
42. Langfelder, P., Zhang, B. & Horvath, S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* **24**, 719–720 (2008).
43. Halkidi, M., Batistakis, Y. & Vazirgiannis, M. On clustering validation techniques. *J. Intell. Inf. Syst.* **17**, 107–145 (2001).
44. Rousseeuw, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).
45. Angerer, P. *et al.* destiny – diffusion maps for large-scale single-cell data in R. *bioRxiv* <http://dx.doi.org/10.1101/023309> (2015).
46. Burnham, K. P. & Anderson, D. R. in *Model Selection and Multimodel Inference A Practical Information-Theoretic Approach* 2nd edn, Ch. 2 (Springer, 2002).
47. Breiman, L. Random Forests. *Mach. Learn.* **45**, 5–32 (2001).
48. Liaw, A. & Wiener, M. Classification and Regression by randomForest. *R News* **2**, 18–22 (2002).
49. Bryja, V., Bonilla, S. & Arenas, E. Derivation of mouse embryonic stem cells. *Nature Protocols* **1**, 2082–2087 (2006).
50. Tanaka, Y. *et al.* Circulation-independent differentiation pathway from extraembryonic mesoderm toward hematopoietic stem cells via hemogenic angioblasts. *Cell Reports* **8**, 31–39 (2014).
51. Sroczynska, P., Lancrin, C., Pearson, S., Kouskoff, V. & Lacaud, G. In vitro differentiation of mouse embryonic stem cells as a model of early hematopoietic development. *Methods Mol. Biol.* **538**, 317–334 (2009).
52. Wilson, N. K. *et al.* Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* **7**, 532–544 (2010).
53. Zwart, W. *et al.* A carrier-assisted ChIP-seq method for estrogen receptor-chromatin interactions from breast cancer core needle biopsy samples. *BMC Genomics* **14**, 232 (2013).
54. Sánchez-Castillo, M. *et al.* CODEX: a next-generation sequencing experiment database for the haematopoietic and embryonic stem cell communities. *Nucleic Acids Res.* **43**, D1117–D1123 (2015).
55. Zhang, Y. *et al.* Model-based analysis of ChIP-seq (MACS). *Genome Biol.* **9**, R137 (2008).
56. Wu, H. & Ji, H. PolyPeak: detecting transcription factor binding sites from ChIP-seq using peak shape information. *PLoS ONE* **9**, e89694 (2014).
57. Wilkinson, D. G. *In Situ Hybridization* (Oxford Univ. Press, 1999).
58. Beddington, R. S. & Robertson, E. J. Axis development and early asymmetry in mammals. *Cell* **96**, 195–209 (1999).
59. Du, J. *et al.* O-fucosylation of thrombospondin type 1 repeats restricts epithelial to mesenchymal transition (EMT) and maintains epiblast pluripotency during mouse gastrulation. *Dev. Biol.* **346**, 25–38 (2010).
60. Donnison, M. *et al.* Loss of the extraembryonic ectoderm in *Elf5* mutants leads to defects in embryonic patterning. *Development* **132**, 2299–2308 (2005).
61. Mitsunaga, K. *et al.* Loss of PGC-specific expression of the orphan nuclear receptor ERR- β results in reduction of germ cell number in mouse embryos. *Mech. Dev.* **121**, 237–246 (2004).
62. Baldwin, H. S. *et al.* Platelet endothelial cell adhesion molecule-1 (PECAM-1/CD31): alternatively spliced, functionally distinct isoforms expressed during mammalian cardiovascular development. *Development* **120**, 2539–2553 (1994).
63. Naiche, L. A., Arora, R., Kania, A., Lewandoski, M. & Papaioannou, V. E. Identity and fate of *Tbx4*-expressing cells reveal developmental cell fate decisions in the allantois, limb, and external genitalia. *Dev. Dyn.* **240**, 2290–2300 (2011).
64. Tamplin, O. J. *et al.* Microarray analysis of *Foxa2* mutant mouse embryos reveals novel gene expression and inductive roles for the gastrula organizer and its derivatives. *BMC Genomics* **9**, 511 (2008).
65. Vincent, S. D. *et al.* *Prdm1* functions in the mesoderm of the second heart field, where it interacts genetically with *Tbx1*, during outflow tract morphogenesis in the mouse embryo. *Hum. Mol. Genet.* **23**, 5087–5101 (2014).
66. Morkel, M. *et al.* β -Catenin regulates Cripto- and Wnt3-dependent gene expression programs in mouse axis and mesoderm formation. *Development* **130**, 6283–6294 (2003).
67. Niwa, H., Miyazaki, J. & Smith, A. G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature Genet.* **24**, 372–376 (2000).
68. Pearce, J. J. H. & Evans, M. J. Mml, a mouse Mix-like gene expressed in the primitive streak. *Mech. Dev.* **87**, 189–192 (1999).
69. Pennisi, D. *et al.* Mutations in Sox18 underlie cardiovascular and hair follicle defects in ragged mice. *Nature Genet.* **24**, 434–437 (2000).
70. Gordon, E. J., Gale, N. W. & Harvey, N. L. Expression of the hyaluronan receptor LYVE-1 is not restricted to the lymphatic vasculature; LYVE-1 is also expressed on embryonic blood vessels. *Dev. Dyn.* **237**, 1901–1909 (2008).
71. Kallianpur, A. R., Jordan, J. E. & Brandt, S. J. The SCL/TAL-1 gene is expressed in progenitors of both the hematopoietic and vascular systems during embryogenesis. *Blood* **83**, 1200–1208 (1994).
72. Robb, L. *et al.* Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene. *Proc. Natl Acad. Sci. USA* **92**, 7075–7079 (1995).
73. Tanaka, Y. *et al.* The transcriptional programme controlled by Runx1 during early embryonic blood development. *Dev. Biol.* **366**, 404–419 (2012).
74. North, T. *et al.* Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development* **126**, 2563–2575 (1999).
75. Palis, J., McGrath, K. E. & Kingsley, P. D. Initiation of hematopoiesis and vasculogenesis in murine yolk sac explants. *Blood* **86**, 156–163 (1995).
76. Drissen, R. *et al.* The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. *Mol. Cell. Biol.* **25**, 5205–5214 (2005).
77. Southwood, C. M., Downs, K. M. & Bieker, J. J. Erythroid Krüppel-like factor exhibits an early and sequentially localized pattern of expression during mammalian erythroid ontogeny. *Dev. Dyn.* **206**, 248–259 (1996).
78. Silver, L. & Palis, J. Initiation of murine embryonic erythropoiesis: a spatial analysis. *Blood* **89**, 1154–1164 (1997).
79. Lanctôt, C., Lamolet, B. & Drouin, J. The bicoid-related homeoprotein Ptx1 defines the most anterior domain of the embryo and differentiates posterior from anterior lateral mesoderm. *Development* **124**, 2807–2817 (1997).
80. Lania, G., Ferrentino, R. & Baldini, A. TBX1 represses Vegfr2 gene expression and enhances the cardiac fate of VEGFR2+ cells. *PLoS ONE* **10**, e0138525 (2015).
81. Brown, C. B. *et al.* Cre-mediated excision of *Fgf8* in the *Tbx1* expression domain reveals a critical role for *Fgf8* in cardiovascular development in the mouse. *Dev. Biol.* **267**, 190–202 (2004).
82. Brennan, J. *et al.* Nodal signalling in the epiblast patterns the early mouse embryo. *Nature* **411**, 965–969 (2001).
83. Meno, C. *et al.* Mouse Lefty2 and zebrafish antiviral are feedback inhibitors of nodal signaling during vertebrate gastrulation. *Mol. Cell* **4**, 287–298 (1999).
84. Bessho, Y. *et al.* Dynamic expression and essential functions of Hes7 in somite segmentation. *Genes Dev.* **15**, 2642–2647 (2001).
85. Oginuma, M., Niwa, Y., Chapman, D. L. & Saga, Y. Mesp2 and Tbx6 cooperatively create periodic patterns coupled with the clock machinery during mouse somitogenesis. *Development* **135**, 2555–2562 (2008).
86. Forlani, S., Lawson, K. A. & Deschamps, J. Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. *Development* **130**, 3807–3819 (2003).
87. Zeigler, B. M. *et al.* The allantois and chorion, when isolated before circulation or chorio-allantoic fusion, have hematopoietic potential. *Development* **133**, 4183–4192 (2006).
88. Downs, K. M., Hellman, E. R., McHugh, J., Barrickman, K. & Inman, K. E. Investigation into a role for the primitive streak in development of the murine allantois. *Development* **131**, 37–55 (2004).
89. Caprioli, A., Jaffredo, T., Gautier, R., Dubourg, C. & Dieterlen-Lièvre, F. Blood-borne seeding by hematopoietic and endothelial precursors from the allantois. *Proc. Natl Acad. Sci. USA* **95**, 1641–1646 (1998).
90. van Nes, J. *et al.* The *Cdx4* mutation affects axial development and reveals an essential role of Cdx genes in the ontogenesis of the placental labyrinth in mice. *Development* **133**, 419–428 (2006).
91. Yang, J. T., Rayburn, H. & Hynes, R. O. Cell adhesion events mediated by α_4 integrins are essential in placental and cardiac development. *Development* **121**, 549–560 (1995).
92. Solloway, M. J. & Robertson, E. J. Early embryonic lethality in *Bmp5;Bmp7* double mutant mice suggests functional redundancy within the 60A subgroup. *Development* **126**, 1753–1768 (1999).
93. Drake, C. J. & Fleming, P. A. Vasculogenesis in the day 6.5 to 9.5 mouse embryo. *Blood* **95**, 1671–1679 (2000).
94. Lee, D. *et al.* ER71 acts downstream of BMP, Notch, and Wnt signaling in blood and vessel progenitor specification. *Cell Stem Cell* **2**, 497–507 (2008).
95. Carapuç, M., Nôvoa, A., Bobola, N. & Mallo, M. Hox genes specify vertebral types in the presomitic mesoderm. *Genes Dev.* **19**, 2116–2121 (2005).
96. Zhang, H. *et al.* Expression of podocalyxin separates the hematopoietic and vascular potentials of mouse embryonic stem cell-derived mesoderm. *Stem Cells* **32**, 191–203 (2014).

97. Herrmann, B. G. Expression pattern of the *Brachyury* gene in whole-mount TWis/TWis mutant embryos. *Development* **113**, 913–917 (1991).
98. Weidgang, C. E. *et al.* TBX3 directs cell-fate decision toward mesendoderm. *Stem Cell Rep.* **1**, 248–265 (2013).
99. Perea-Gómez, A., Shawlot, W., Sasaki, H., Behringer, R. R. & Ang, S. *HNf3β* and *Lim1* interact in the visceral endoderm to regulate primitive streak formation and anterior-posterior polarity in the mouse embryo. *Development* **126**, 4499–4511 (1999).
100. Saga, Y. *et al.* MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. *Development* **126**, 3437–3447 (1999).
101. Trimborn, T., Gribnau, J., Grosveld, F. & Fraser, P. Mechanisms of developmental control of transcription in the murine α - and β -globin loci. *Genes Dev.* **13**, 112–124 (1999).
102. Kingsley, P. D., Malik, J., Fantauzzo, K. A. & Palis, J. Yolk sac-derived primitive erythroblasts enucleate during mammalian embryogenesis. *Blood* **104**, 19–25 (2004).
103. Hodge, D. *et al.* A global role for EKLF in definitive and primitive erythropoiesis. *Blood* **107**, 3359–3370 (2006).
104. Isern, J. *et al.* Single-lineage transcriptome analysis reveals key regulatory pathways in primitive erythroid progenitors in the mouse embryo. *Blood* **117**, 4924–4934 (2011).
105. Joshi, A., Hannah, R., Diamanti, E. & Göttgens, B. Gene set control analysis predicts hematopoietic control mechanisms from genome-wide transcription factor binding data. *Exp. Hematol.* **41**, 354–366.e14 (2013).
106. Goode, D. K. *et al.* Dynamic gene regulatory networks drive hematopoietic specification and differentiation. *Dev. Cell* **36**, 572–587 (2016).



Extended Data Figure 1 | FACS of single cells. **a**, Flk1⁺ cells were sorted from three embryos at primitive streak and head fold stages and four embryos at neural plate stage. Labels such as 'S4' refer to the embryo number in the metadata available online at <http://gastrulation.stemcells.cam.ac.uk/scialdone2016>. **b**, CD41⁺Flk1⁻ cells (red gate) and CD41⁺Flk1⁺

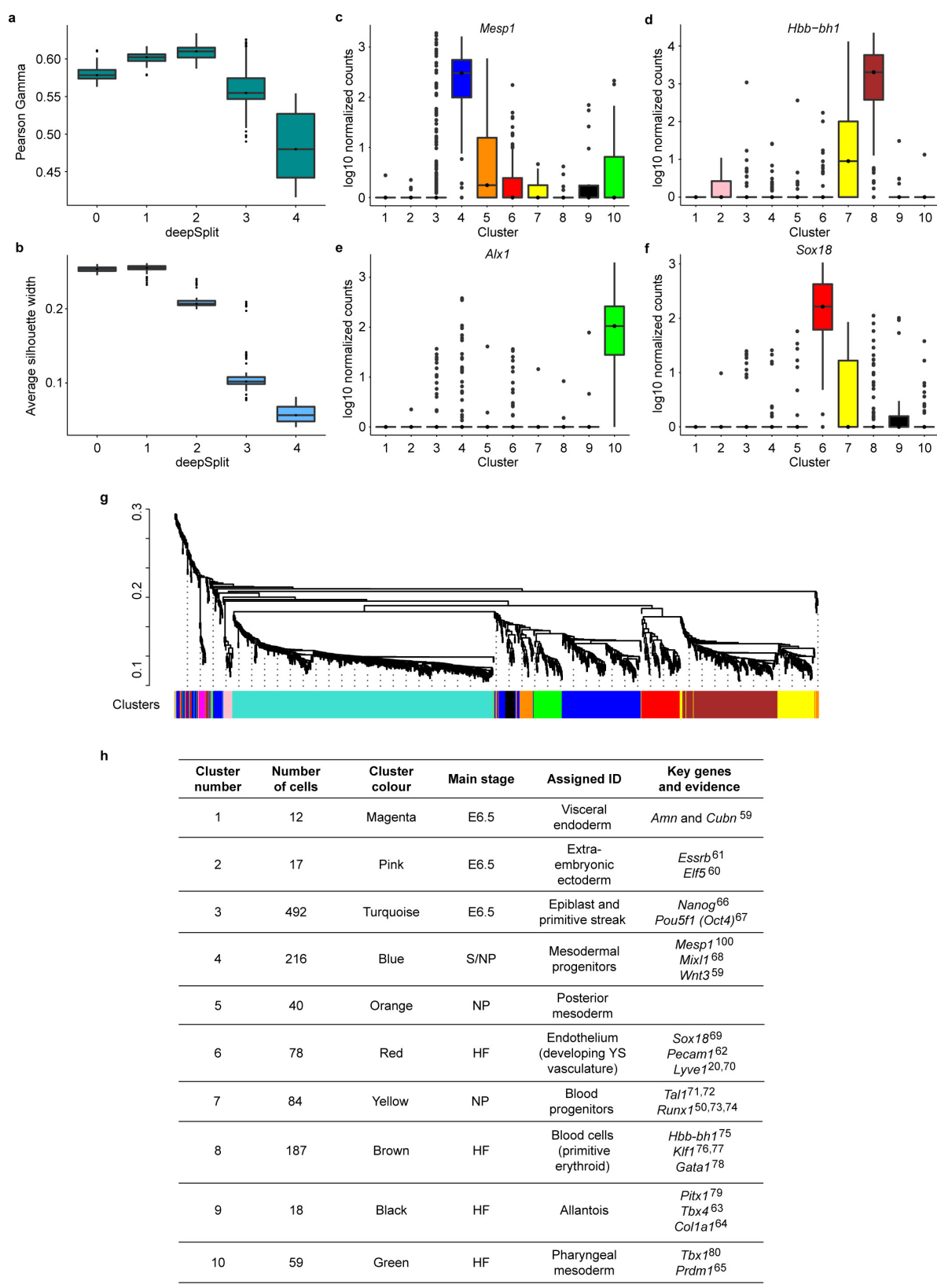
(green gate) cells were sorted from eight embryos each at neural plate and head fold stages (see Fig. 1 for cell numbers). Each stage was sorted on two occasions. Labels above FACS plots refer to the sort and embryo number in the metadata available online, as above. In all plots, pink text indicates the percentage of cells in that gate.



Extended Data Figure 2 | Quality control of single-cell RNA-seq data.

a, Table showing numbers and estimates of numbers of cells of different phenotypes present in embryos between E6.5 and E7.75 (head fold stage) and numbers sorted for this study. *Total cell numbers for E6.5 are from Beddington and Robertson (1999)⁵⁸. †Total numbers and numbers of Flk1⁺ cells are from Moignard *et al.*, (2015)¹⁵. Percentages of cells expressing Flk1 and/or CD41 at neural plate and head fold stages are the average values from the embryos used in this study and were used to calculate the estimated numbers present in embryos from the total cell numbers. ND, not done. **b**, t-SNE representation of the five metrics used

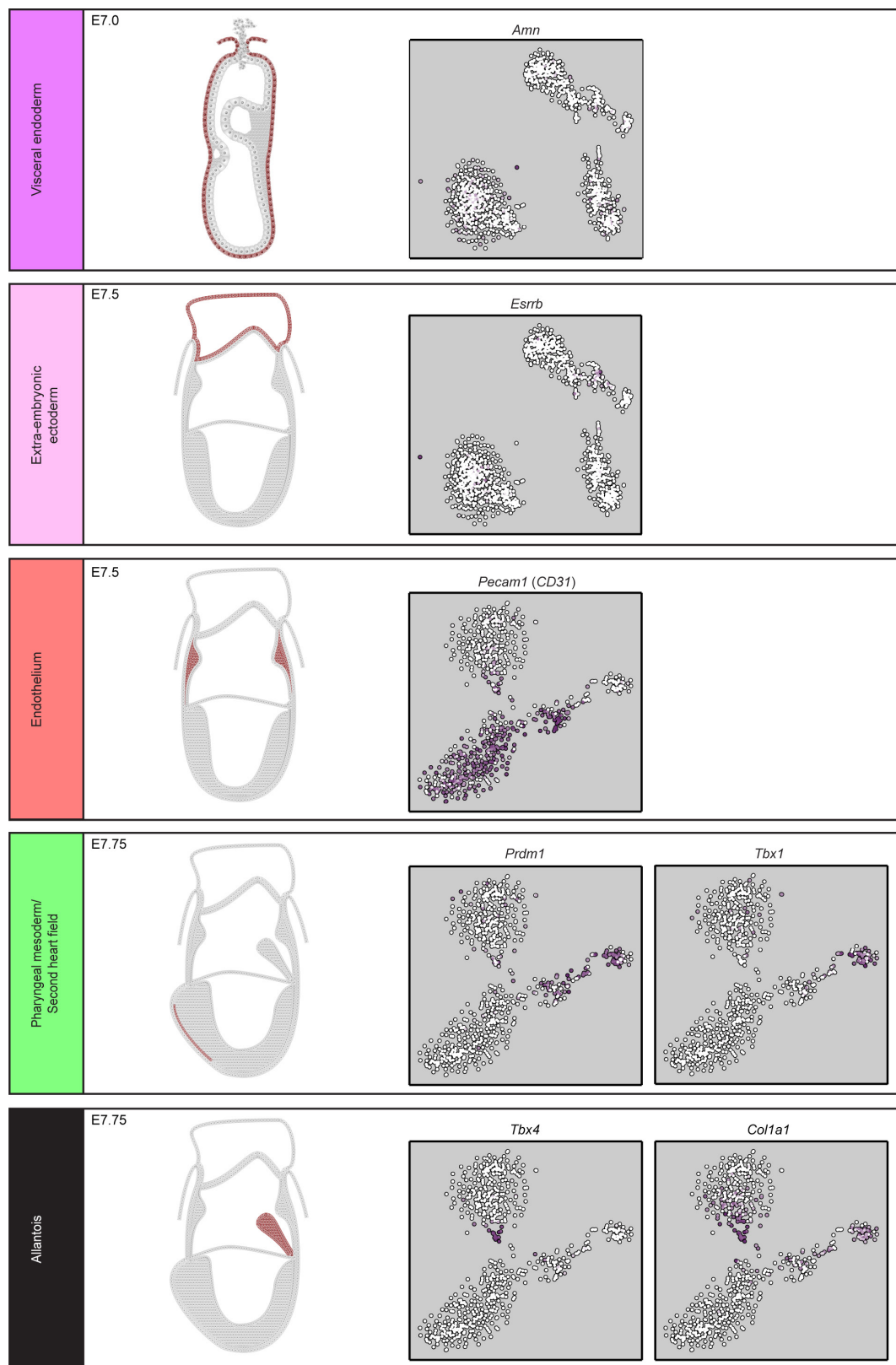
to assess the quality of all 2,208 sorted cells from the wild-type and *Tal1* experiments. Only cells that passed all criteria (blue circles) were used for downstream analysis. **c**, Squared coefficient of variation (CV^2) as a function of the mean normalized counts (μ) for genes across all cells. The green line shows the fit $CV^2 = a_1/\mu + \alpha_0$. All highly variable genes (with an adjusted P value < 0.1) are marked by red circles. **d**, Number of genes detected (that is, with more than ten normalized read counts) in cells across the different clusters in the WT (left) and the *Tal1*^{-/-} (right) mice. Boxes indicate the median and interquartile range.



Extended Data Figure 3 | See next page for caption.

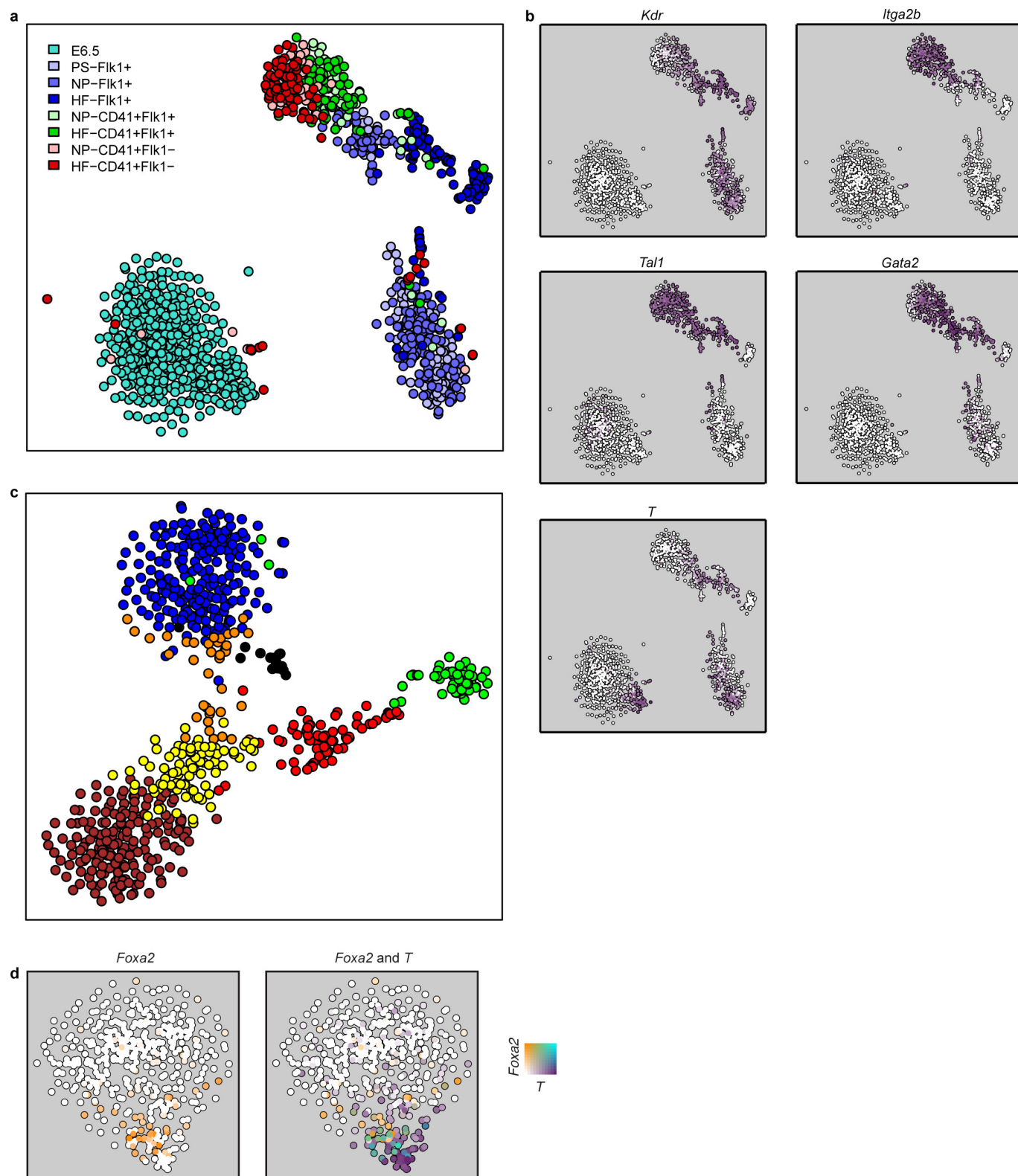
Extended Data Figure 3 | Identifying cell clusters. The dynamic hybrid cut algorithm was used with all possible values of the 'deepSplit' parameter on 100 bootstrapped subsamples. **a, b**, To assess the quality of the clustering, the Pearson gamma (**a**) and the average silhouette width (**b**) were calculated. Higher values of these parameters correspond to better clustering. The Pearson gamma represents the correlation between the dissimilarity of samples and a binary variable that equals 0 for pairs of samples in the same cluster and 1 for samples in different clusters. The average silhouette width measures the average separation between neighbouring clusters^{43,44}. At 'deepSplit' = 2 the Pearson gamma is highest whereas the average silhouette width begins to decrease. This suggests that at such a value of the 'deepSplit' parameter a good compromise between robustness and sensitivity is achieved. The Pearson gamma and

the average silhouette width were computed with the R function 'cluster.stats' in the 'fpc' package (version 2.1-9). **c–f**, Examples of marker genes for four clusters: *Mesp1* for cluster 4 (top-ranked marker) (**c**), *Hbb-bh1* for cluster 8 (fourth-ranked) (**d**), *Alx1* for cluster 10 (top-ranked) (**e**) and *Sox18* for cluster 6 (second-ranked) (**f**). The y axis shows the log₁₀-normalized expression of the genes. For **a–f**, boxes indicate the median and interquartile range. **g**, Dendrogram showing the clustering of the cells in the first experiment. The colours at the bottom indicate the cluster each cell was assigned to by the dynamic hybrid cut algorithm. Cluster assignment was used to sort cells in Fig. 1b. **h**, Identities were assigned to the ten clusters in Fig. 1c on the basis of the expression of key genes^{20,50,59–80} associated with various mesodermal lineages or spatial locations within the embryo.



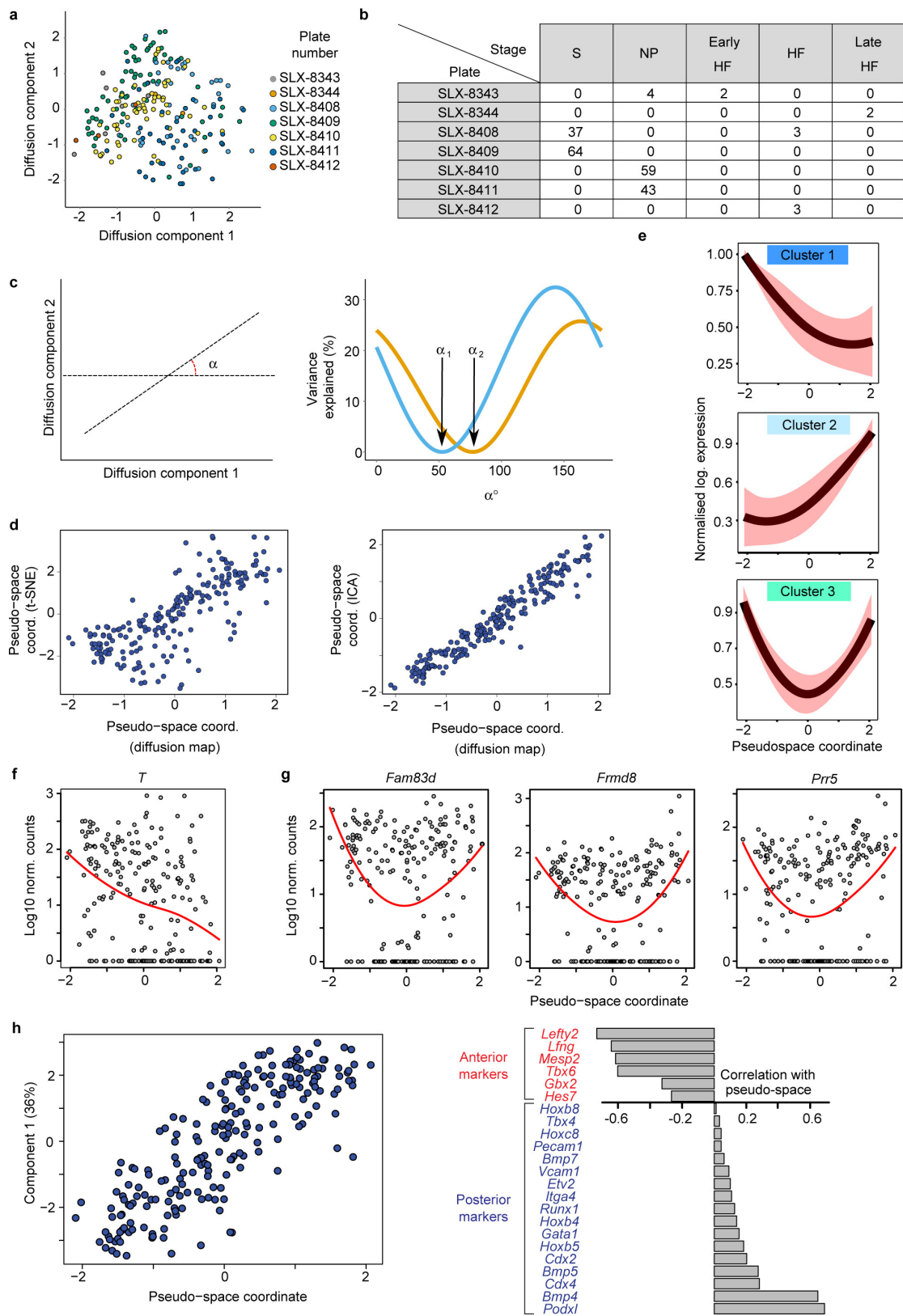
Extended Data Figure 4 | Expression of key marker genes in E7.0–7.75 embryos. Schematic representations of expression patterns were generated from published *in situ* hybridization data (see citations) for key markers of clusters 1 (magenta, visceral endoderm⁵⁹), 2 (pink, extra-embryonic ectoderm⁶¹), 6 (red, yolk sac endothelium⁶²), 9 (black, allantois^{63,64})

and 10 (green, second heart field^{65,81}). Anterior is shown on the left and posterior on the right. Also shown is the t-SNE for all 1,205 cells or 682 cells from E7.0 onwards (primitive streak, neural plate and head fold stages) indicating expression of each gene (white, low; purple, high).



Extended Data Figure 5 | Expression of key genes used for sorting single cells. a, t-SNE as in Fig. 1 showing the sorting strategy for each of the 1,205 cells. **b**, Expression of Flk1 (*Kdr*), CD41 (*Itga2b*), Scl (*Tal1*), *Gata2* and *T* (Brachyury) superimposed onto the t-SNE. **c**, t-SNE showing only

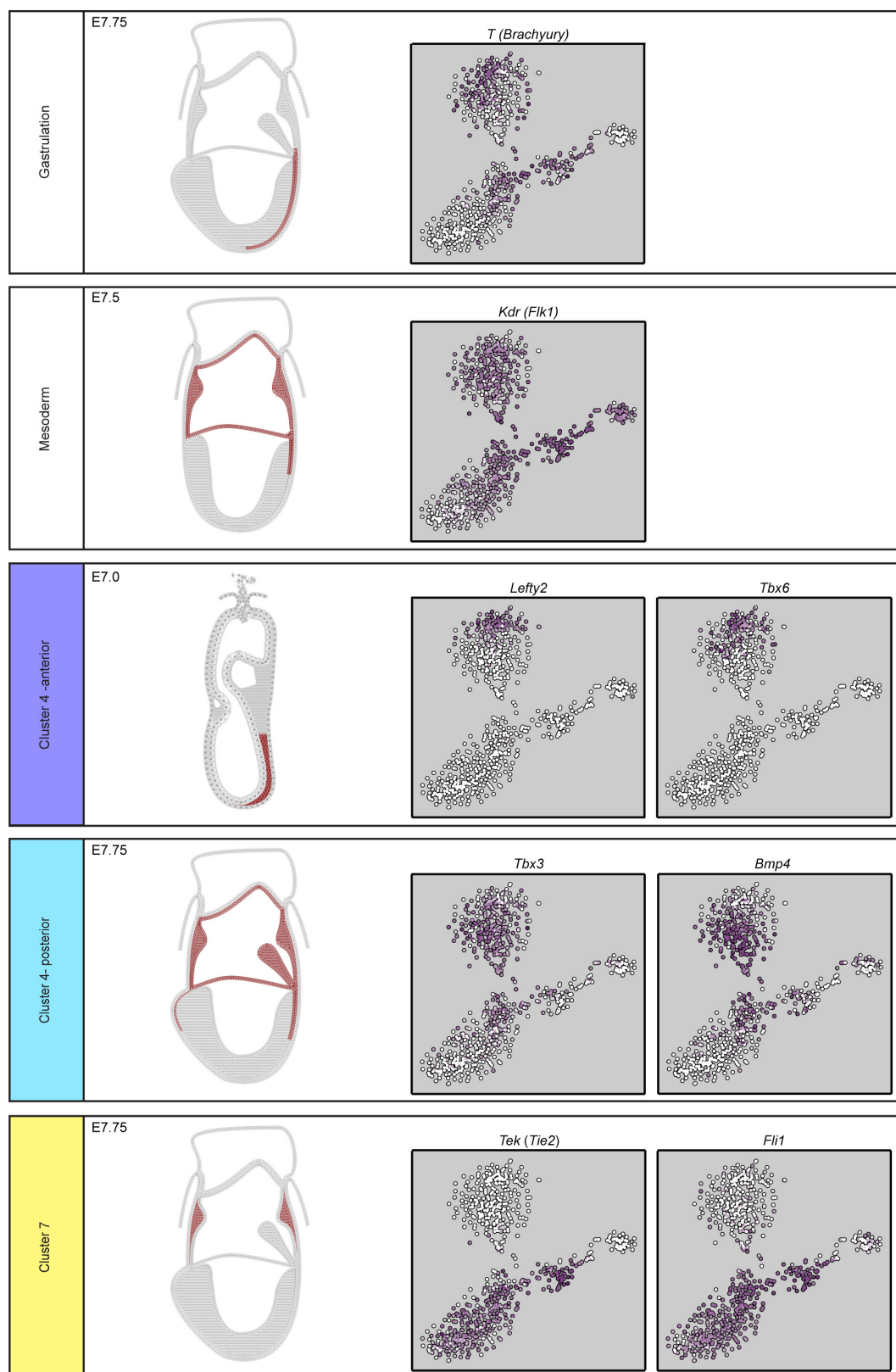
the 682 cells from primitive streak, neural plate and head fold stages, coloured according to cluster as in Fig. 1c. **d**, t-SNE for the 481 E6.5 cells in cluster 3, as in Fig. 2a. Each point is coloured by expression of *T* and *Foxa2*.



Extended Data Figure 6 | See next page for caption.

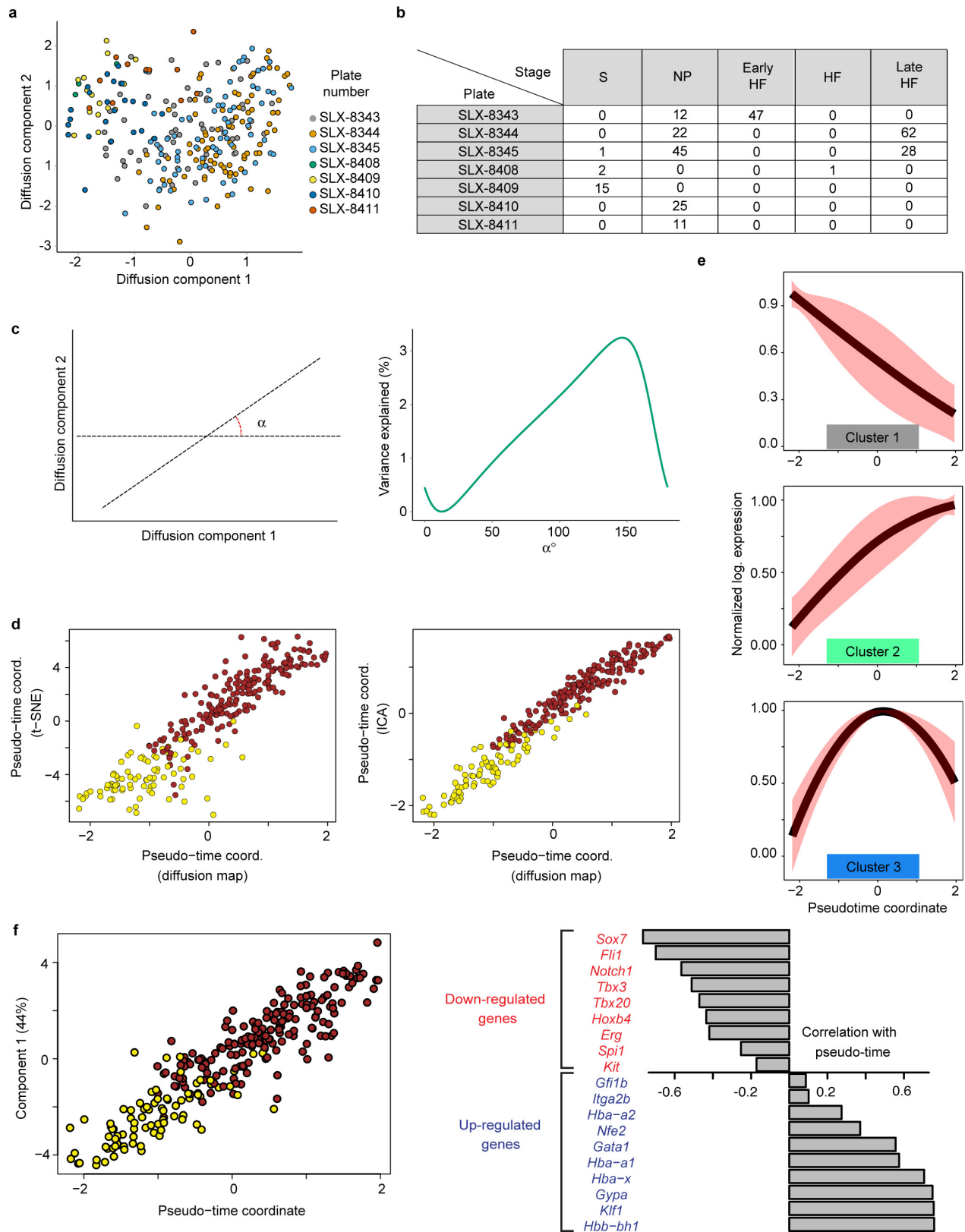
Extended Data Figure 6 | Pseudospace analysis of cluster 4 correlates with anterior–posterior position along the primitive streak. **a**, Diffusion plot of the 216 cells in cluster 4. Different colours correspond to different plates and different lanes of flow cells. **b**, Table showing the number of cells in each stage analysed on the different lanes of flow cells (S, primitive streak; NP, neural plate; HF, head fold). **c**, A direction in the diffusion space can be identified by the angle α that it forms with the first diffusion component (left panel). For each value of α the right panel shows the percentage of variance explained by the batch effect associated to plates SLX-8408 and SLX-8409 (orange line) and plates SLX-8410 and SLX-8411 (blue line). Labels α_1 and α_2 are the angles corresponding to directions that correlate the least with the batch effect (that is, variance explained by the batch effect is minimum). **d**, The use of alternative dimensionality reduction techniques results in the identification of highly correlated pseudospace coordinates. A t-SNE projection of the dissimilarity matrix was performed (perplexity set to 50), and the direction corresponding to the pseudospace coordinate was estimated by minimizing the correlation with the batch effect (left panel; Spearman correlation between the two pseudospace coordinates 0.79, $P < 2.2 \times 10^{-16}$). Independent component analysis was performed on the dissimilarity matrix with the ‘fastICA’ R function, and three independent components (corresponding to the two batch effects and the biological

effect) were estimated. The presumptive pseudospace coordinate is the component having the smallest correlation with the batch effects (right panel; Spearman correlation coefficient is 0.97, $P < 2.2 \times 10^{-16}$). **e**, Plots showing the average expression of genes in clusters 1–3 of Fig. 3c along the pseudospace axis. Gene expression levels are normalized between 0 and 1. Dark red lines indicate the normalized mean expression levels of genes in each cluster as obtained from the fitting procedure and red shaded area indicates standard deviation. **f**, Expression of *T* as function of the pseudospace coordinate. **g**, Gene expression levels for example genes showing high-low-high expression pattern across the blue cluster. In **f** and **g**, putative anterior cells are to the left and posterior to the right. Each dot represents a cell and red lines indicate fits based on local polynomial functions (see Methods). **h**, We performed principal component analysis on the cells in cluster 4 by using markers of pre-somitic mesoderm as anterior mesoderm markers and genes expressed in haemato-vascular and allantoic mesoderm as posterior markers^{82–95}, as well as *Podxl* which was shown to separate distinct Flk1⁺ mesodermal lineages⁹⁶. The first component explained 36% of the total variance and was highly correlated with the pseudospace coordinate (left; Spearman rank correlation 0.84, $P < 2.2 \times 10^{-16}$). All the anterior markers were negatively correlated with the pseudospace coordinate, whereas all posterior markers had a positive correlation (right).



Extended Data Figure 7 | Expression of key genes along the anterior–posterior axis of the primitive streak in E7.0–7.75 embryos. Schematic representations of gene expression were generated from published *in situ* hybridization data (see citations) for key markers of clusters 4 (blue, mesoderm) and 7 (yellow, posterior mesoderm/blood progenitors). Expression of *T* (Brachyury)⁹⁷ and *Flk1* (Kdr, from in-house data) are shown to illustrate the extent of the primitive streak at E7.5. *Lefty2* and *Tbx6* (ref. 59) are expressed in the putative anterior portion of cluster 4 and in more anterior regions of the primitive streak in *in situ* analysis.

Tbx3 (ref. 98) and *Bmp4* (ref. 99) are expressed in the more posterior portion of cluster 4 and in the embryo are expressed in the more posterior region of the primitive streak around the amnion and into the extra-embryonic mesoderm. *Tek* and *Fli1* (from in-house data) are expressed in cluster 7 and in the embryo are found exclusively in the extra-embryonic portion. Also shown is the t-SNE for the cells from E7.0 onwards (primitive streak, neural plate and head fold stages) indicating expression of each gene (white, low; purple, high).

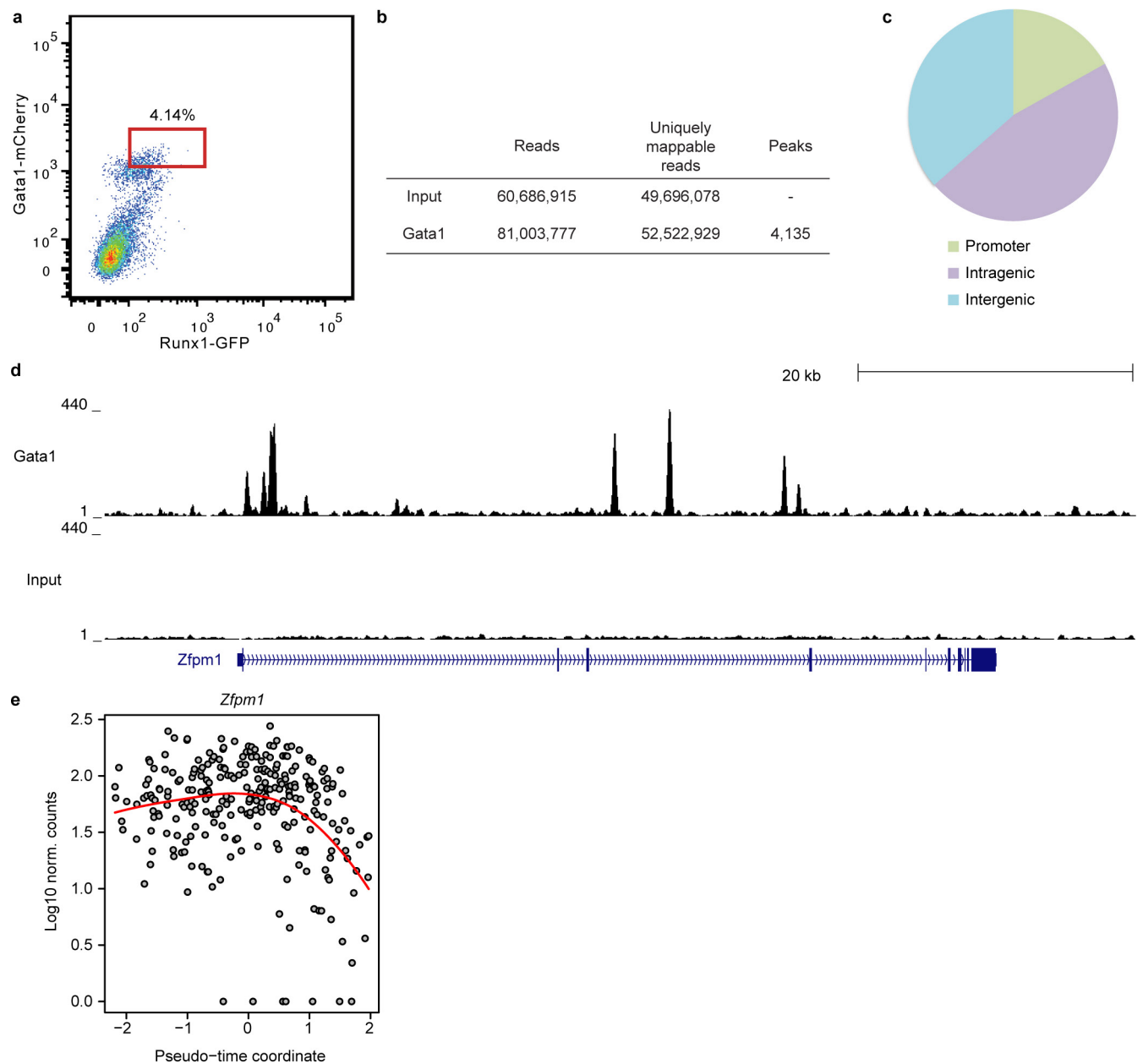


Extended Data Figure 8 | See next page for caption.

Extended Data Figure 8 | Pseudotime analysis of primitive erythroid development.

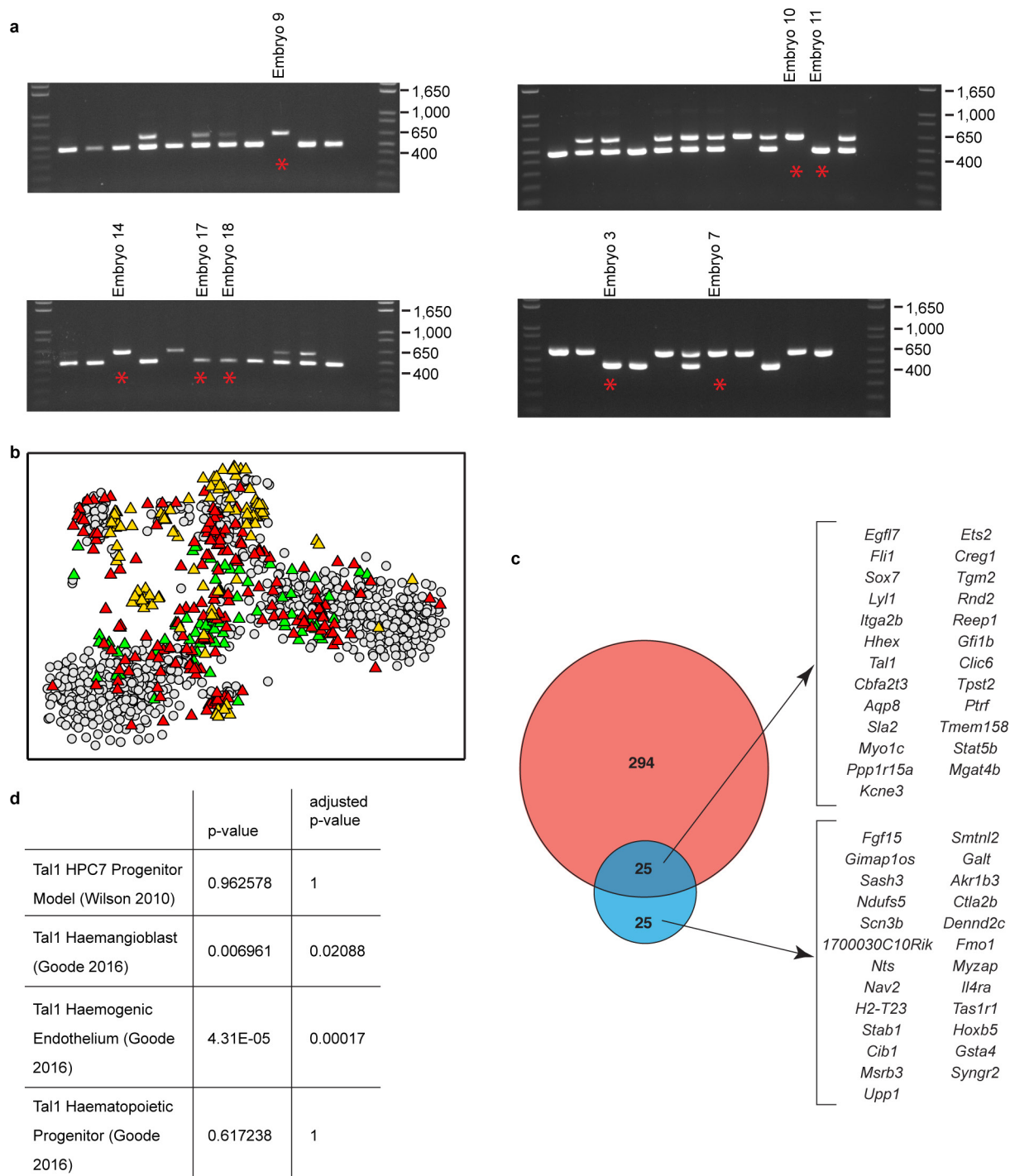
a, Diffusion plot of the 271 cells in clusters 7 and 8. Different colours correspond to different plates and lanes of flow cells. **b**, Table showing the number of cells in each stage collected on the different plates (S, primitive streak; NP, neural plate; HF, head fold). **c**, Analogously to Extended Data Fig. 6, the angle α identifies a direction in the diffusion space (left panel). The percentage of variance explained by the batch effect associated to plates SLX-8344 and SLX-8345 is plotted as a function of α in the right panel. **d**, The pseudotime coordinate is robust to the use of different dimensionality reduction techniques, as shown in the left panel with t-SNE (Spearman correlation 0.92, $P < 2.2 \times 10^{-16}$) and in the right panel with independent component analysis (Spearman correlation 0.97, $P < 2.2 \times 10^{-16}$; same procedure described in Extended

Data Fig. 6d). **e**, Plots showing the average expression of genes in clusters 1–3 of Fig. 4c along the pseudotime axis. Gene expression levels are normalized between 0 and 1. Dark red lines are the average expression levels of genes in each cluster as obtained from the fitting procedure, after normalization. Red shaded areas indicate standard deviation. **f**, Principal component analysis was performed on the expression pattern of genes known from previous studies to be upregulated or downregulated along the blood developmental trajectory^{15,66,100–104}. The first principal component (explaining 44% of total variance) showed a very strong correlation with the pseudotime coordinate (left; Spearman correlation coefficient 0.91, $P < 2.2 \times 10^{-16}$). All upregulated (downregulated) genes positively (negatively) correlate with the pseudotime coordinate (right).



Extended Data Figure 9 | ChIP-seq for Gata1 in ESC-derived haematopoietic cells. **a**, Flow cytometry for Gata1-mCherry and Runx1-IRES-GFP knock-in reporter genes in embryoid body cells after 5 days of haematopoietic differentiation. Cells were sorted for the expression of both Runx1-IRES-GFP and Gata1-mCherry knock-in reporter genes to provide *in vitro* equivalents of the developing primitive erythrocytes assayed by RNA-seq. The gate used for sorting is shown in red. **b**, Numbers of reads and peaks identified for Gata1 and an input sample after mapping

and peak calling; 4,135 Gata1 peaks were identified. **c**, Distribution of Gata1 peaks between promoter, intragenic and intergenic sequences. **d**, University of California, Santa Cruz Genome Browser tracks for Gata1 and input sample at the *Zfp1* (*Fog1*) locus known to be a target of Gata1, indicating the quality of the ChIP-seq data. **e**, Expression of Gata1 target *Zfp1* during the pseudotimecourse for erythroid development, as in Fig. 4.



Extended Data Figure 10 | Collection of embryos from *Tal1 LacZ*^{+/+} crosses. **a**, Genotyping PCR for embryos from *Tal1 LacZ*^{+/+} crosses. Lower band is the WT allele and upper band is the mutant allele carrying a neomycin knock in. Presence of both bands indicates heterozygosity. Embryos from which sequencing data were obtained are indicated with a red star and the number given corresponds to embryo identity in the metadata available online with the sequencing data. **b**, t-SNE as in Fig. 5d showing *Tal1* data (triangles; 377 cells) and original WT data (grey circles; 1,205 cells). *Tal1* data are coloured according to the embryo stage

from which they were collected: green, neural plate; red, head fold; orange, four-somite pair. **c**, As in Fig. 5d, showing the complete list of genes. **d**, Gene set control analysis¹⁰⁵ was used to identify statistically significant overlaps between genes significantly downregulated in *Tal1*^{-/-} compared with WT cells in the endothelial cluster (see Fig. 5) and *Tal1* targets identified by ChIP-seq. Gene set control analysis identified an enrichment of our gene set with *Tal1* ChIP-seq in ESC-derived haemangioblasts and haemogenic endothelium¹⁰⁶, but not in ESC-derived haematopoietic progenitors¹⁰⁶ or a haematopoietic progenitor cell line⁵².

Cloche is a bHLH-PAS transcription factor that drives haemato-vascular specification

Sven Reischauer^{1,2*}, Oliver A. Stone^{1,2*}, Alethia Villasenor^{1,2*}, Neil Chi^{1,3}, Suk-Won Jin^{1†}, Marcel Martin⁴, Miler T. Lee^{5†}, Nana Fukuda², Michele Marass², Alec Witty³, Ian Fiddes^{1†}, Taiyi Kuo^{1†}, Won-Suk Chung^{1†}, Sherveen Salek^{1†}, Robert Lerrigo^{1†}, Jessica Alsiö^{1†}, Shujun Luo^{6†}, Dominika Tworus⁷, Sruthy M. Augustine², Sophie Mucenicks², Björn Nystedt⁸, Antonio J. Giraldez⁵, Gary P. Schroth⁶, Olov Andersson⁷ & Didier Y. R. Stainier^{1,2}

Vascular and haematopoietic cells organize into specialized tissues during early embryogenesis to supply essential nutrients to all organs and thus play critical roles in development and disease. At the top of the haemato-vascular specification cascade lies *cloche*, a gene that when mutated in zebrafish leads to the striking phenotype of loss of most endothelial and haematopoietic cells^{1–4} and a significant increase in cardiomyocyte numbers⁵. Although this mutant has been analysed extensively to investigate mesoderm diversification and differentiation^{1–7} and continues to be broadly used as a unique avascular model, the isolation of the *cloche* gene has been challenging due to its telomeric location. Here we used a deletion allele of *cloche* to identify several new *cloche* candidate genes within this genomic region, and systematically genome-edited each candidate. Through this comprehensive interrogation, we succeeded in isolating the *cloche* gene and discovered that it encodes a PAS-domain-containing bHLH transcription factor, and that it is expressed in a highly specific spatiotemporal pattern starting during late gastrulation. Gain-of-function experiments show that it can potently induce endothelial gene expression. Epistasis experiments reveal that it functions upstream of *etv2* and *tal1*, the earliest expressed endothelial and haematopoietic transcription factor genes identified to date. A mammalian *cloche* orthologue can also rescue blood vessel formation in zebrafish *cloche* mutants, indicating a highly conserved role in vertebrate vasculogenesis and haematopoiesis. The identification of this master regulator of endothelial and haematopoietic fate enhances our understanding of early mesoderm diversification and may lead to improved protocols for the generation of endothelial and haematopoietic cells *in vivo* and *in vitro*.

Endothelial cells are specialized epithelial cells that line the lumen of blood vessels, and have been the focus of many studies investigating how the vasculature develops and functions^{8–10}. More recent studies have revealed that a subset of endothelial cells, known as haemogenic endothelium, can give rise to haematopoietic stem cells (HSCs) that eventually generate the complete range of adult blood cells¹¹. A number of transcription factors including *Etv2* (also known as *Etsrp*), *Tal1* (also known as *Scl*), *Sox7*, as well as other *Ets*, *SoxF*, *Gata* and *Forkhead* family members have been implicated in the control of endothelial and blood cell differentiation^{12–15}; however, it remains unclear how

endothelial and blood cells are initially specified from the mesoderm. The zebrafish *cloche* mutant, which lacks most endothelial (Fig. 1a) and haematopoietic cells but exhibits an expanded heart field and increased cardiomyocyte numbers⁵, constitutes a unique tool to investigate the diversification of the mesoderm into the cardiovascular and haematopoietic lineages. On the basis of these findings and because these mesodermal cell types are hypothesized to derive from a common cardiovascular precursor^{16–19}, it has been proposed that *cloche* functions at the time when the mesoderm is specified into endothelial and haematopoietic progenitors, acting at the top of the haemato-vascular specification cascade. As a result, both the mutant and the identity of the gene have been the subject of intense investigation for more than two decades^{1–7}. However, the isolation of the gene has been challenging due to its telomeric location⁴ and the lack of a reliable genome assembly at chromosome ends. Through CRISPR/Cas9 mutagenesis, complementation studies, identification of a severe molecular lesion in an ENU (*N*-ethyl-*N*-nitrosourea)-induced allele, knockdown experiments and rescue data, here we report the isolation of the *cloche* gene and show that it encodes a PAS (PER-ARNT-SIM)-domain-containing bHLH transcription factor. We further show, through gain- and loss-of-function analyses, that *Cloche* functions at the top of the endothelial and haematopoietic hierarchy, and that this function is probably conserved throughout vertebrate evolution.

We reported previously the mapping of the *cloche* locus to one of the telomeres of chromosome 13, and the identification of a tightly linked CA repeat marker, z1496, based on genotyping 2,359 mutant embryos⁴. To provide higher genetic resolution to the *cloche* region and narrow down the interval containing the *cloche* mutation, we genotyped an additional 7,920 mutant embryos for a total of 10,279 embryos, which allowed us to identify additional zero recombinant markers, as well as closely linked flanking markers (Fig. 1b and Supplementary Table 1). Searching the available genomic databases led to the identification of several genes in the *cloche* locus including *lclat1* (also known as *lycat*, previously reported to be a strong candidate for the *cloche* gene)²⁰, *lbh*, *klhdc3*, and *inpp5a*. We sequenced these genes in several ENU-induced *cloche* alleles looking for severe genetic alterations, and also carried out functional studies including morpholino knockdowns and mRNA rescues (data not shown), but failed to generate fully convincing data that any of them corresponded to the *cloche* gene. Because the physical

¹Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, California, USA. ²Max Planck Institute for Heart and Lung Research, Department of Developmental Genetics, Bad Nauheim 61231, Germany. ³Department of Medicine, Division of Cardiology, Institute of Genomic Medicine, University of California San Diego, La Jolla, California 92037, USA. ⁴Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, Solna 17121, Sweden. ⁵Department of Genetics, Yale University School of Medicine, New Haven, Connecticut 06520, USA. ⁶illumina, San Diego, California 92122, USA. ⁷Department of Cell and Molecular Biology, Karolinska Institutet, Stockholm 17177, Sweden. ⁸Science for Life Laboratory, Department of Cell and Molecular Biology, Uppsala University, Uppsala 75124, Sweden. [†]Present addresses: School of Life Sciences, Gwangju Institute of Science and Technology, Gwangju, 61005, South Korea and Yale Cardiovascular Research Center, Yale University School of Medicine, New Haven, Connecticut 06511, USA (S.-W.J.); Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA (M.T.L.); Genomics Institute, University of California Santa Cruz and Howard Hughes Medical Institute, Santa Cruz, California 95064, USA (I.F.); Department of Medicine and Berrie Diabetes Center, Columbia University College of Physicians and Surgeons, New York, New York 10032, USA (T.K.); Department of Biological Sciences, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, South Korea (W.-S.C.); Johns Hopkins Hospital, Wilmer Eye Institute, Baltimore, Maryland 21224, USA (S.S.); Division of General Internal Medicine, University of Washington, Seattle, Washington 98104, USA (R.L.); Novartis, Basel 4056, Switzerland (J.A.); Personalis, Menlo Park, California 94025, USA (S.L.).

*These authors contributed equally to this work.

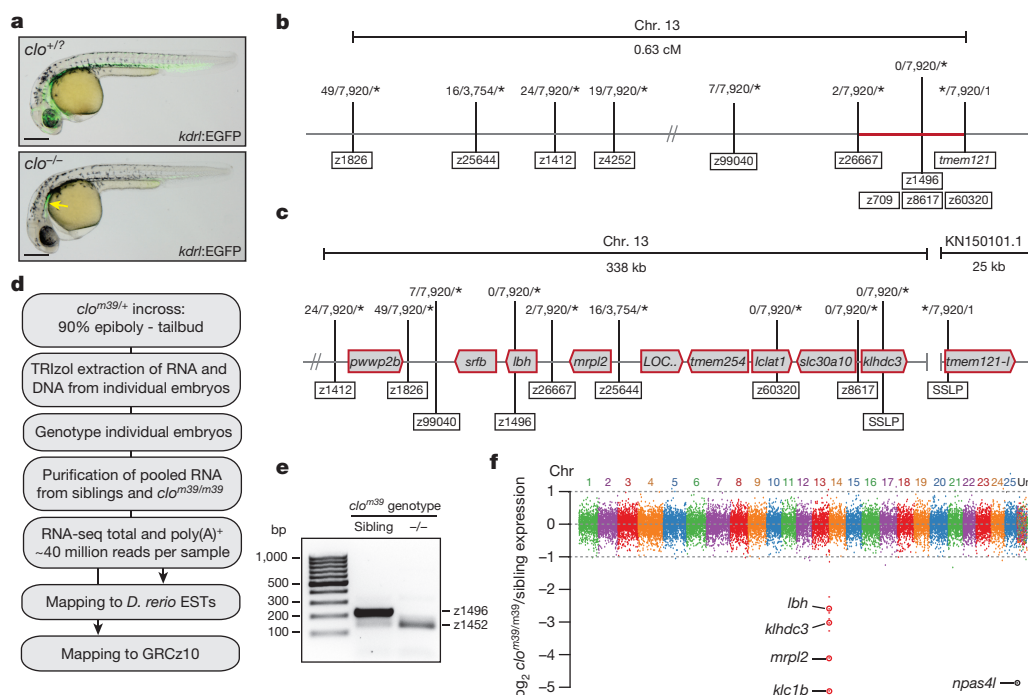


Figure 1 | Identification and analysis of candidate genes from the *cloche* locus. **a**, A 30 hours post-fertilization (hpf) sibling and *clo*^{s5} mutant; *kdr:EGFP* expression marks endothelial cells, which are mostly missing in *cloche* mutants; yellow arrow, *kdr:EGFP* expression in pharyngeal arch endoderm observed in both wild-type and *clo* mutant embryos. Scale bars, 200 μ m. **b**, Genetic map of *cloche* locus showing meiotic recombination events; critical region in red. **c**, Physical map of *cloche* locus based on

GRCz10, including 338 kb sub-telomeric region with more than 25 gaps and unassigned scaffold KN150101.1; meiotic recombination data superimposed. **d**, Strategy to identify novel *clo* candidates using *clo*^{m39} deletion allele. **e**, Genotyping using *z1496* (deleted in *clo*^{m39/m39}) and *z1452* (present in *clo*^{m39/m39}). **f**, RNA-seq expression differences between *clo*^{m39/m39} and wild-type siblings; genes sorted by chromosomal position.

assembly of the *cloche* region, which harbours a large number of gaps, did not match our genetic map (Fig. 1c), we reasoned that additional genes were located within the region and that one of them could be *cloche*. To identify these missing genes, we used the *cloche* deletion allele *m39*, which lacks a large number of genetic markers around the *cloche* locus². *In situ* hybridization analyses examining the expression of the earliest known regulators of endothelial and haematopoietic cell differentiation, the ETS domain protein gene *etv2* (refs 6, 12) and the bHLH protein gene *tal1* (refs 3, 13–15) revealed that the earliest time differences could be observed between *cloche* mutants and their wild-type siblings was during early somitogenesis^{3,7} (data not shown). We thus collected a large number of single embryos from *clo*^{m39} heterozygous intercrosses between 90% epiboly and tail bud stages, as the *cloche* gene should be expressed before the onset of the phenotype. We extracted both DNA and RNA from these individual embryos, and after genotyping them with the simple sequence repeat *z1496*, which is deleted in *clo*^{m39} mutants, and *z1452*, which is present in *clo*^{m39} mutants (Fig. 1d, e), pooled wild-type and mutant RNA. Sequencing total and polyA-enriched RNA generated ~40 million reads per sample, which were compared to the entire collection of *Danio rerio* expressed sequence tags (ESTs), and also mapped to GRCz10, the latest zebrafish genome assembly at the time. This bioinformatics analysis led to the identification of several new genes, as well as genes previously identified to be in the *cloche* region (Fig. 1f and Supplementary Table 2), all of which were then mutated using TALEN or CRISPR/Cas9 technology. For each of these genes, identified founders carrying probable deleterious mutations were crossed to the *clo*^{m39} allele for complementation testing, and once available, the F1 fish were intercrossed to assess vascular morphology by examining *Tg(kdr:EGFP)* expression in F2 embryos. Out of the 19 mutants tested in this manner (Extended Data Fig. 1), only mutants for a gene showing homology to *npas4* failed to complement the *clo*^{m39} allele, and displayed a vascular phenotype in the F2 generation. Confirming our speculation, this gene did not map

to the physical assembly of the *cloche* region but to an unassigned 25 kb scaffold (KN150101.1; Fig. 1c). From the RNA-seq data, we had identified a sequence translating into a continuous reading frame of 293 amino acids that showed 33% identity over a length of 113 amino acids to the C-terminal region of Npas4 from *Cynoglossus semilaevis*, and used it to design a CRISPR guide RNA (gRNA) to induce a molecular lesion in the gene. To isolate a full-length cDNA, we extended the sequence of this fragment using 5' and 3' rapid amplification of cDNA ends (RACE) (Extended Data Fig. 2) and assembled a clone of 4,286 base pairs (bp) encoding a PAS-domain-containing bHLH protein of 647 amino acids. This protein shares highest, though limited, homology with NPAS4 (Extended Data Fig. 3), which in mouse has been reported to play a key role in activity-regulated neuronal gene expression and memory formation^{21,22}. NPAS4 is a member of the class I bHLH-PAS family of proteins which includes the aryl hydrocarbon receptor (AHR)²³, as well as the hypoxia inducible factors HIF1A²⁴ and HIF2A²⁵, three proteins implicated in vascular and blood development in vertebrates. The zebrafish genome comprises two well-annotated *npas4* genes: *npas4a* and *npas4b*. On the basis of our analysis of synteny conservation and protein similarity (Extended Data Fig. 3), we will hereafter refer to this gene as *npas4l*, although it should be noted that the corresponding protein's primary structure is substantially divergent from that of Npas4a and Npas4b, especially in its C terminus. In order to assemble the *npas4l* genomic region, and as the KN150101.1 scaffold did not contain the full *npas4l* sequence, we used whole-genome PacBio and Illumina sequencing data to extend it from a length of 22 kilobases (kb) to 35 kb. Mapped RNA-seq reads and gene predictions by AUGUSTUS show that *npas4l*, like other *npas4* genes, consists of eight exons (Fig. 2a), extending from 24 kb to 29 kb on the improved scaffold. In addition, the coding sequence for *klhdc3* could be aligned in part to both a region upstream of *npas4l* and the end of the GRCz10 chromosome 13 sequence, confirming that our extended version of KN150101.1 should be placed

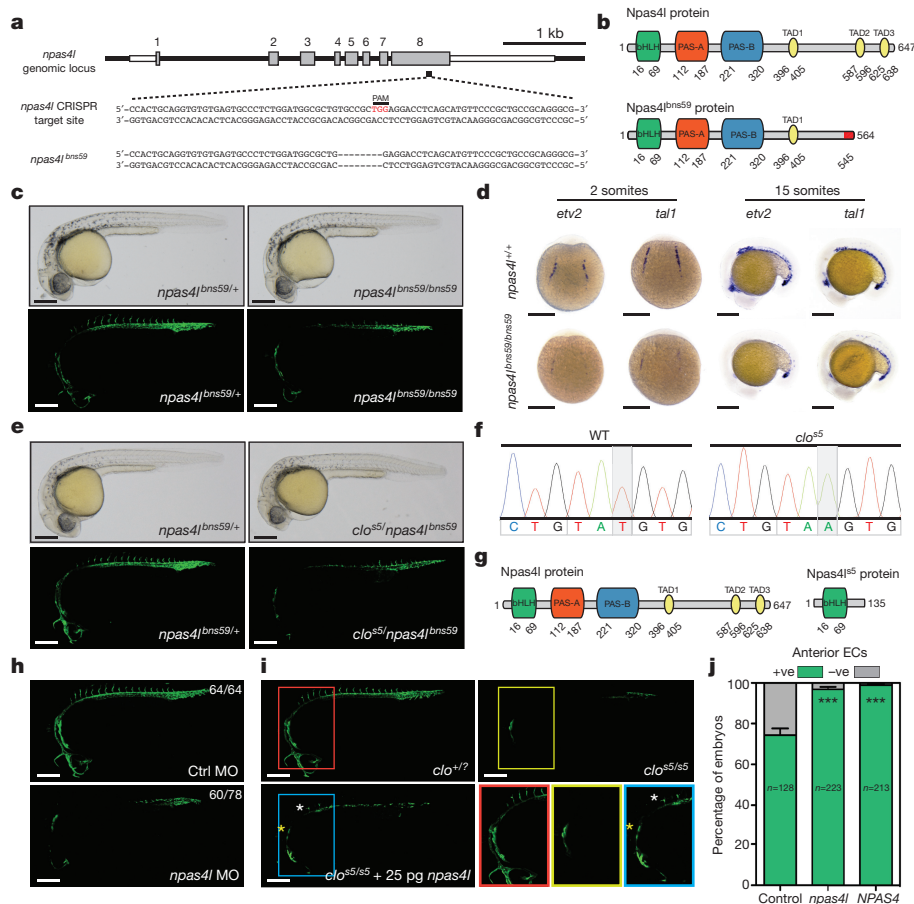


Figure 2 | *npas4l* is *cloche*. **a**, *npas4l* locus, CRISPR target site, and *npas4l*^{bns59} allele. **b**, Npas4l and Npas4l^{bns59} proteins. The *bns59* mutation causes a premature stop codon at amino acid 564 and loss of 2 out of 3 putative transactivation domains (TADs). **c**, Brightfield and confocal projections of 28 hpf *Tg(kdrl:EGFP)* sibling and *npas4l*^{bns59/bns59}. The *npas4l*^{bns59/bns59} *kdrl:EGFP* phenotype is clearly milder than *clo*^{m39/m39} or *clo*^{s5/s5}, as expected from the underlying molecular lesion. **d**, Whole-mount *in situ* hybridization for *etv2* and *tal1* expression in *npas4l*^{+/+} and *npas4l*^{bns59/bns59}. **e**, *npas4l*^{bns59} fails to complement *clo*^{s5}; 28 hpf *Tg(kdrl:EGFP)* embryos. **f**, Single-base mutation affecting Y136 leads to premature stop codon in *clo*^{s5}. **g**, Npas4l and Npas4l^{s5} proteins. **h**, Confocal projections of 26 hpf *Tg(kdrl:EGFP)* embryos injected with control or *npas4l* morpholinos; *x/y*, number of embryos showing representative

phenotype (*x*), number of embryos examined (*y*). **i**, Confocal projections of 28 hpf *Tg(kdrl:EGFP)* uninjected *clo*^{+/+} and *clo*^{s5/s5} embryos, and a *clo*^{s5/s5} embryo injected with 25 pg *npas4l* mRNA; high magnification views of boxed areas shown in lower-right panel; yellow asterisk, anterior endothelium; white asterisk, anterior intersegmental vessel. The *Tg(kdrl:EGFP)* expression in pharyngeal arch endoderm was observed in all embryos. **j**, Quantification of mRNA rescue of *clo*^{s5/s5}; 25 pg of zebrafish *npas4l* or human NPAS4 mRNA injected into one-cell stage embryos from *clo*^{s5} intercrosses; presence of anterior endothelial cells based on *Tg(kdrl:EGFP)* expression was recorded in all embryos at 28 hpf. Values represent means \pm s.e.m. ****P* < 0.001 by one-way analysis of variance (ANOVA) followed by Tukey's HSD test. Scale bars, 200 μ m.

in reverse-complementary orientation downstream of *klhdc3* on chromosome 13 (Extended Data Fig. 2).

To test whether *npas4l* is indeed the *cloche* gene, we first analysed the phenotype of the CRISPR-generated allele, *npas4l*^{bns59}, which carries an 8 bp deletion in the 3' end of the gene leading to a premature stop codon at amino acid 564 after 19 missense amino acids (Fig. 2b). These mutants display fully penetrant vascular and blood defects (Fig. 2c, d and Extended Data Fig. 4), clearly recapitulating the *cloche* phenotype^{1–3}; however, these defects appear weaker than those observed in other *cloche* alleles (Extended Data Fig. 4), consistent with the fact that the predicted Npas4l^{bns59} protein only lacks two of the three putative transactivation domains in its C terminus (Fig. 2b). Most notably, *npas4l*^{bns59} mutants develop some endocardial cells (Extended Data Fig. 4), which are never observed in *clo*^{m39/m39} or *clo*^{s5/s5}, an ENU-induced allele²⁶. Complementation analyses between *npas4l*^{bns59} and *clo*^{s5} showed that these two mutations also fail to complement (Fig. 2e), indicating that they are allelic. Sequencing the *npas4l* cDNA from *clo*^{s5} mutants revealed a stop codon at position 135, within the first PAS domain (Fig. 2f, g). We also injected into wild-type embryos a combination of ATG and splice morpholinos for *npas4l* (Extended

Data Fig. 5), and observed similarly severe blood vessel development defects as those exhibited by *clo*^{s5} mutants (Fig. 2h). Finally, we found that injection of 25 pg of zebrafish *npas4l* or human NPAS4 mRNA into embryos from *clo*^{s5} heterozygous intercrosses rescued the anterior population of *Tg(kdrl:EGFP)*-expressing endothelial cells, as well as the formation of intersegmental vessels in *clo*^{s5} mutants (Fig. 2i, j and Extended Data Fig. 6). Thus, the results from the genetic mapping and complementation studies with a CRISPR-generated allele, the identification of a severe molecular lesion in an ENU-induced allele, and data from morpholino knockdowns and mRNA rescues, together strongly indicate that *npas4l* is the *cloche* gene.

We next examined the spatiotemporal pattern of *npas4l* expression in comparison with that of *etv2* and *tal1*. Quantitative reverse transcription PCR (RT-qPCR) experiments showed that *npas4l* expression peaks before the end of gastrulation and is only weakly expressed by 24 h post-fertilization (hpf) (Extended Data Fig. 7), the stage when the heart starts beating, whereas *etv2* and *tal1* expression appear to peak at the mid-somitogenesis stages (Extended Data Fig. 7). Analysis of published RNA-seq data sets²⁷ confirmed these observations (Extended Data Fig. 7). Whole-mount *in situ* hybridization studies show that *npas4l*

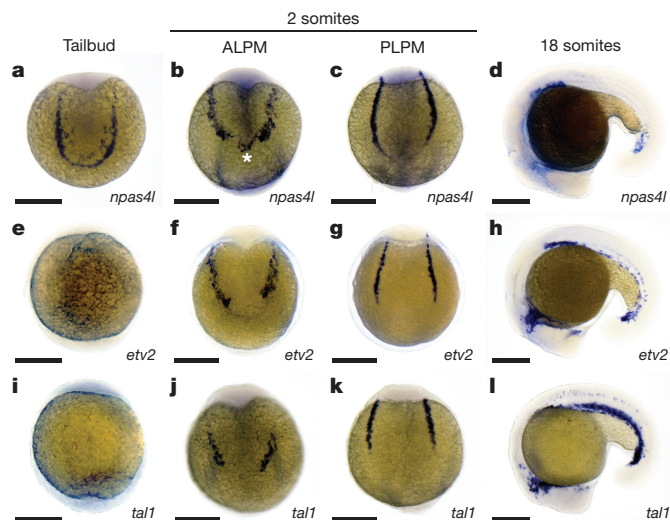


Figure 3 | *cloche/npas4l* is expressed in endothelial and haematopoietic precursors, and this expression precedes that of the earliest endothelial and haematopoietic markers *etv2* and *tal1*. a–l, Whole-mount *in situ* hybridization for *npas4l*, *etv2* and *tal1* expression; posterior part of tailbud stage embryos shown in a, e and i. ALPM, anterior lateral plate mesoderm; PLPM, posterior lateral plate mesoderm. Asterisk (b) marks rostromedial population of *npas4l*⁺ cells that at this stage are *etv2* and *tal1* negative. Scale bars, 200 μ m.

expression precedes the onset of *etv2* and *tal1* expression (Fig. 3a, e, i), and that all three genes display a very similar expression pattern in the anterior and posterior lateral plate mesoderm, the tissue containing the endothelial and haematopoietic precursors. This expression pattern in the posterior lateral plate mesoderm is also similar to that exhibited by *gata1a*, which marks erythroid lineage cells and is clearly visible by the two-somite stage (Extended Data Fig. 7). Notably, *npas4l* expression at early stages is uniquely observed in a rostromedial population of cells in the anterior lateral plate mesoderm (Fig. 3b; asterisk) that is initially negative for *etv2* and *tal1* expression (Fig. 3f, j), and probably corresponds to myeloid precursors. Similarly, in the posterior lateral plate mesoderm, *npas4l* expression appears to extend more posteriorly (Fig. 3c) compared to *etv2* and *tal1* expression (Fig. 3g, k), which may reflect its earlier activation in these endothelial and haematopoietic precursors. By the 18-somite stage, the levels of *npas4l* expression are clearly decreasing (Fig. 3d), consistent with the RT–qPCR data, while the levels of *etv2* and *tal1* expression remain high (Fig. 3h, l). Thus, the onset of *clo/npas4l* expression precedes that of *etv2* and *tal1*; *clo/npas4l* expression is also observed in cells that appear to be initially negative for *etv2* or *tal1* expression, and after peaking at late gastrulation/early somitogenesis stages, it is rapidly downregulated.

To position *clo/npas4l* in the hierarchy of endothelial and haematopoietic regulators, we conducted gain- and loss-of-function experiments with *etv2* and *tal1*. Injection of *etv2* or *tal1* mRNA into one-cell stage embryos had no effect on *npas4l* expression when assessed by RT–qPCR (Fig. 4a). Similarly, morpholino-mediated knock-down of *etv2* or *tal1* (Extended Data Fig. 5) had no effect on *npas4l* expression when assessed by RT–qPCR (Fig. 4a) or whole-mount *in situ* hybridizations (Fig. 4b). In contrast, injection of *npas4l* mRNA into one-cell stage embryos led to a broad and rapid upregulation of *etv2* (Fig. 4c) and *tal1* (Fig. 4d) expression, as well as *lmo2*, *fli1a*, *fli1b*, *sox7* and *sox18* (Extended Data Fig. 8) expression. Together, these data show that *Clo/Npas4l* regulates the expression of the earliest known regulators of endothelial and haematopoietic differentiation, placing it at the top of the hierarchy (Fig. 4e). However, injections of *npas4l* mRNA into one-cell stage embryos did not induce *gata1a* expression (data not shown). Similarly, injection of *npas4l* mRNA into *clo*⁵⁵ mutants did not consistently rescue the *gata1a* expression defect, consistent with previous observations²⁸ that *Cloche* functions cell non-autonomously

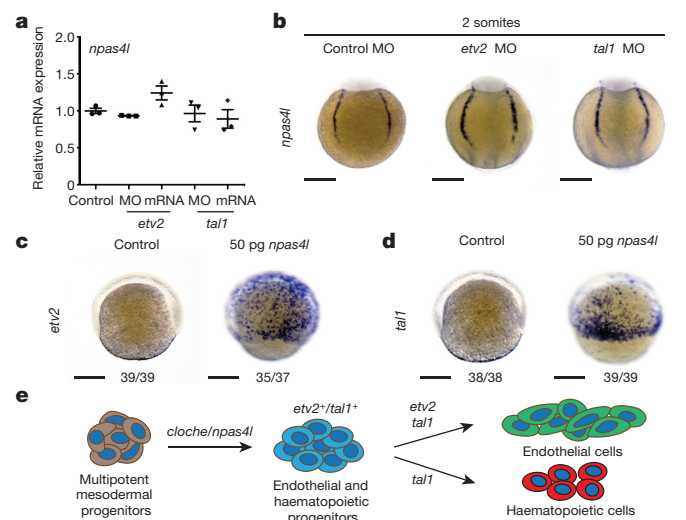


Figure 4 | *Cloche/Npas4l* is a master regulator of the endothelial differentiation cascade. a, RT–qPCR analysis of *npas4l* expression in tailbud stage embryos injected with *etv2* morpholino (MO), *etv2* mRNA, *tal1* MO or *tal1* mRNA at the one-cell stage, normalized to uninjected controls. Error bars, s.e.m.; $n = 3$ biological replicates; no significant changes observed. b, Whole-mount *in situ* hybridization for *npas4l* expression in control, *etv2* and *tal1* morphants ($n > 40$ for each MO). c, d, Whole-mount *in situ* hybridization for *etv2* (c) and *tal1* (d) expression at shield stage in control and *npas4l* mRNA injected wild-type embryos; x/y , number of embryos showing representative phenotype (x), number of embryos examined (y). e, Schematic illustration of *Cloche/Npas4l* function; *cloche/npas4l* is necessary and sufficient for the expression of early endothelial and haematopoietic markers including *etv2* and *tal1*. Scale bars, 200 μ m.

in erythropoiesis at a step before *gata1* expression. It is therefore possible that *Cloche/Npas4l* function needs to be downregulated in red blood cell precursors for their differentiation, and further experiments will be needed to test this hypothesis.

The identification of the *Cloche* protein as a transcriptional regulator has allowed further analysis of its function. Extensive loss-of-function, and now gain-of-function, studies clearly show that *Cloche* functions at the top of a transcriptional cascade that drives the commitment of multipotent mesodermal cells to the endothelial and haematopoietic lineages. Considering its highly restricted expression pattern, it is probable that investigating the extrinsic signals that induce *clo/npas4l* expression will lead to additional insights into the control of mesoderm diversification, including the commitment to the endothelial and haematopoietic lineages, and away from the cardiomyocyte lineage. Similarly, the identification of the *Cloche* protein will facilitate the discovery of additional regulators of endothelial and haematopoietic development. Although the *cloche/npas4l* gene is clearly present from lampreys to birds (Extended Data Fig. 3), the identification of its functional orthologue(s) in mammals will require further, comprehensive investigation. To this end, we have observed that *NPAS4* is transiently expressed in human pluripotent stem cells during their differentiation into endothelial cells (Extended Data Fig. 9); however, given that *Npas4* mutant mice survive to adulthood and thus are unlikely to have severe vascular defects, we speculate that other class I bHLH-PAS proteins may function redundantly with *NPAS4* to regulate mesoderm diversification in mammals. Consistent with this notion, we have observed that the expression of *AHR*, another class I bHLH-PAS transcription factor gene, becomes activated as *NPAS4* is downregulated in these human pluripotent stem cells differentiating into endothelial cells (Extended Data Fig. 9). Supporting the recent finding that *AHR* may modulate the expansion and differentiation of human haematopoietic progenitors²³, we have found that human *AHR* can induce the expression of *Tal1* and *Lmo2* in mouse P19 cells (Extended Data Fig. 10).

Class I bHLH-PAS proteins have to dimerize with class II bHLH-PAS proteins in order to function, and recent genetic analyses have shown that mouse embryos lacking the class II bHLH-PAS protein ARNT contain a substantially reduced number of angioblasts²⁹. Similarly, embryonic stem cell differentiation studies show that loss of ARNT disrupts endothelial specification, while activation of class I bHLH-PAS proteins by hypoxia promotes it³⁰. Notably, ARNT function appears to be VEGF independent during angioblast specification³⁰. Together, these data strongly implicate members of the bHLH-PAS protein family in controlling the specification of the haemato-vascular system throughout evolution, and call for a detailed functional analysis of mammalian bHLH-PAS proteins in this process. More broadly, the identification of *Cloche* and analysis of its function should provide new avenues towards developing novel strategies for lineage reprogramming of stem/progenitor cells into endothelium and blood for regenerative medicine.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 27 December 2015; accepted 25 May 2016.

1. Stainier, D. Y., Weinstein, B. M., Detrich, H. W., III, Zon, L. I. & Fishman, M. C. *cloche*, an early acting zebrafish gene, is required by both the endothelial and hematopoietic lineages. *Development* **121**, 3141–3150 (1995).
2. Liao, W. *et al.* The zebrafish gene *cloche* acts upstream of a *flk-1* homologue to regulate endothelial cell differentiation. *Development* **124**, 381–389 (1997).
3. Liao, E. C. *et al.* SCL/Tal-1 transcription factor acts downstream of *cloche* to specify hematopoietic and vascular progenitors in zebrafish. *Genes Dev.* **12**, 621–626 (1998).
4. Liao, W., Ho, C. Y., Yan, Y. L., Postlethwait, J. & Stainier, D. Y. Hhex and Scl function in parallel to regulate early endothelial and blood differentiation in zebrafish. *Development* **127**, 4303–4313 (2000).
5. Schoenebeck, J. J., Keegan, B. R. & Yelon, D. Vessel and blood specification override cardiac potential in anterior mesoderm. *Dev. Cell* **13**, 254–267 (2007).
6. Sumanas, S., Jorntak, T. & Lin, S. Identification of novel vascular endothelial-specific genes by the microarray analysis of the zebrafish *cloche* mutants. *Blood* **106**, 534–541 (2005).
7. Liu, F., Walmsley, M., Rodaway, A. & Patient, R. Flt1 acts at the top of the transcriptional network driving blood and endothelial development. *Curr. Biol.* **18**, 1234–1240 (2008).
8. Rossant, J. & Howard, L. Signaling pathways in vascular development. *Annu. Rev. Cell Dev. Biol.* **18**, 541–573 (2002).
9. Adams, R. H. & Alitalo, K. Molecular regulation of angiogenesis and lymphangiogenesis. *Nature Rev. Mol. Cell Biol.* **8**, 464–478 (2007).
10. Herbert, S. P. & Stainier, D. Y. Molecular control of endothelial cell behaviour during blood vessel morphogenesis. *Nat. Rev. Mol. Cell Biol.* **12**, 551–564 (2011).
11. Hirschi, K. K. Hemogenic endothelium during development and beyond. *Blood* **119**, 4823–4827 (2012).
12. De Val, S. & Black, B. L. Transcriptional control of endothelial cell development. *Dev. Cell* **16**, 180–195 (2009).
13. Begley, C. G. *et al.* The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. *Proc. Natl Acad. Sci. USA* **86**, 10128–10132 (1989).
14. Endoh, M., Ogawa, M., Orkin, S. & Nishikawa, S. SCL/*tal-1*-dependent process determines a competence to select the definitive hematopoietic lineage prior to endothelial differentiation. *EMBO J.* **21**, 6700–6708 (2002).
15. Van Handel, B. *et al.* Scl represses cardiomyogenesis in prospective hemogenic endothelium and endocardium. *Cell* **150**, 590–605 (2012).
16. Lacaud, G., Robertson, S., Palis, J., Kennedy, M. & Keller, G. Regulation of hemangioblast development. *Ann. NY Acad. Sci.* **938**, 96–108 (2001).
17. Huber, T. L., Kouskoff, V., Fehling, H. J., Palis, J. & Keller, G. Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. *Nature* **432**, 625–630 (2004).
18. Vogeli, K. M., Jin, S. W., Martin, G. R. & Stainier, D. Y. A common progenitor for haematopoietic and endothelial lineages in the zebrafish gastrula. *Nature* **443**, 337–339 (2006).
19. Yang, L. *et al.* Human cardiovascular progenitor cells develop from a KDR⁺ embryonic-stem-cell-derived population. *Nature* **453**, 524–528 (2008).
20. Xiong, J. W., Yu, Q., Zhang, J. & Mably, J. D. An acyltransferase controls the generation of hematopoietic and endothelial lineages in zebrafish. *Circ. Res.* **102**, 1057–1064 (2008).
21. Ooe, N., Saito, K., Mikami, N., Nakatuka, I. & Kaneko, H. Identification of a novel basic helix-loop-helix-PAS factor, NXF, reveals a Sim2 competitive, positive regulatory role in dendritic-cytoskeleton modulator Drebrin gene expression. *Mol. Cell. Biol.* **24**, 608–616 (2004).
22. Lin, Y. *et al.* Activity-dependent regulation of inhibitory synapse development by Npas4. *Nature* **455**, 1198–1204 (2008).
23. Smith, B. W. *et al.* The aryl hydrocarbon receptor directs hematopoietic progenitor cell expansion and differentiation. *Blood* **122**, 376–385 (2013).
24. Iyer, N. V. *et al.* Cellular and developmental control of O₂ homeostasis by hypoxia-inducible factor 1 α . *Genes Dev.* **12**, 149–162 (1998).
25. Peng, J., Zhang, L., Drysdale, L. & Fong, G. H. The transcription factor EPAS-1/hypoxia-inducible factor 2 α plays an important role in vascular remodeling. *Proc. Natl Acad. Sci. USA* **97**, 8386–8391 (2000).
26. Field, H. A., Dong, P. D., Beis, D. & Stainier, D. Y. Formation of the digestive system in zebrafish. II. Pancreas morphogenesis. *Dev. Biol.* **261**, 197–208 (2003).
27. Pauli, A. *et al.* Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Res.* **22**, 577–591 (2012).
28. Parker, L. & Stainier, D. Y. Cell-autonomous and non-autonomous requirements for the zebrafish gene *cloche* in hematopoiesis. *Development* **126**, 2643–2651 (1999).
29. Ramírez-Bergeron, D. L., Runge, A., Adelman, D. M., Gohil, M. & Simon, M. C. HIF-dependent hematopoietic factors regulate the development of the embryonic vasculature. *Dev. Cell* **11**, 81–92 (2006).
30. Han, Y., Kuang, S. Z., Gomer, A. & Ramírez-Bergeron, D. L. Hypoxia influences the vascular expansion and differentiation of embryonic stem cell cultures through the temporal expression of vascular endothelial growth factor receptors in an ARNT-dependent manner. *Stem Cells* **28**, 799–809 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank all laboratory members who have worked on *cloche* over the years starting with W. Liao and H. Sawyer (UCSF), as well as J. Collins (Sanger Institute) for help in expanding the GRCz10 assembly, W. Coppieters (Liège), Z. Wang (JGI), X. Chen (BGI), H. Yuan (BGI) for their hard work in trying to resolve the *cloche* locus, C. Helker, M. Higuchi and C. Gerri for reagents, discussions and reading of the manuscript, A. Borchers (Marburg) for help with *Xenopus* experiments, and funding from the DFG (S.R.), AHA (S.R., S.-W.J., N.C.), NIH (N.C., A.J.G., D.Y.R.S.), the Ragnar Söderberg Foundation and Swedish Research Council (O.A.), and the Packard Foundation and Max Planck Society (D.Y.R.S.).

Author Contributions S.R., O.S., A.V. and D.Y.R.S. designed and performed experiments identifying *cloche* and its function. N.F., Mi.M., S.M.A. and S.M. performed reverse genetics and downstream analysis. N.C., S.-W.J., T.K., W.-S.C., S.S., R.L. and J.A. worked on the meiotic mapping. S.L., Ma.M., M.T.L., I.F. and D.T. contributed to sequencing and bioinformatics. Human embryonic stem cell assays were handled by A.W. Supervision by D.Y.R.S., N.C., O.A., G.P.S., A.J.G. and B.N. All authors contributed to data analysis and manuscript preparation.

Author Information Zebrafish *npas4l* sequences have been deposited to GenBank (KX066018, KX066019), RNA-seq data deposited to GEO (GSE76690), and PacBio whole-genome sequencing data to SRA (LT571435). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.Y.R.S. (didier.stainier@mpi-bn.mpg.de).

Reviewer Information *Nature* thanks K. Alitalo and K. Poss and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

No statistical methods were used to predetermine sample size. The investigators were inherently blinded by the experimental design during the experiments and outcome assessment for the data shown in Figs 2 (c–e, i, j) and 4 (c, d) and Extended Data Figs 4 and 6, as well as the genotyping for the RNA-seq profiling. The experiments were not randomized.

Zebrafish handling. All zebrafish husbandry was performed under standard conditions in accordance with institutional (UCSF and MPG) and national ethical and animal welfare guidelines. Mutant lines were generated in the *Tg(kdrl:EGFP)^{s843}* background³¹. Complementation tests were performed with *clo^{s5}* and *clo^{m39}* mutant lines. Overexpression studies were performed in AB embryos.

Genetic mapping of the *clo^{s5}* allele and RNA-seq analysis of *clo^{m39}*. For genetic mapping, fine mapping of the *clo^{s5}* allele was performed by genotyping homozygous mutant embryos with simple sequence-length polymorphisms (SSLPs) identified in the locus. Individual genotyping primer sequences can be found in Supplementary Table 1. For RNA-seq analysis, individual embryos from *clo^{m39}* heterozygous intercrosses were collected between 90% epiboly and tailbud. TRIzol extraction was performed on individual embryos to yield DNA and RNA. Briefly, single embryos were homogenized in TRIzol (Invitrogen) and phases were separated with chloroform. For DNA extraction, DNA was precipitated from the organic phase with ethanol, washed twice with 0.1 M sodium citrate in 10% ethanol and once with 75% ethanol. The SSLP markers z1496 (0/7,920) and z1452 (chr. 16) were used for genotyping. Embryos positive for the z1452 PCR product but negative for z1496 were considered homozygous for the *clo^{m39}* deletion. Embryos positive for both products were considered wild-type siblings. For RNA extraction, the aqueous phase extracted from sibling or *clo^{m39/m39}* embryos was pooled separately. RNA was precipitated using isopropanol and washed with 70% ethanol. Genomic DNA was removed by DNase treatment and total RNA was isolated using the Zymo Research RNA Clean & Concentrator kit. RNA-seq libraries were prepared from random primed total RNA, which was normalized using the DSN method to decrease the abundance of high copy transcripts^{32,33}, or from poly(A) selected mRNA using the TruSeq RNA v2 kit from Illumina (RS-122-2001) with an insert size of 200 bp following the manufacturer's recommendations. Single read sequencing was performed on an Illumina Genome Analyzer IIx instrument with 50 bp read lengths. More than 40 million reads were collected for each of the four samples.

RNA-seq data processing. Data from poly(A)⁺ and total RNA sequencing were treated separately. Sequencing reads were aligned non-strand specifically to the zebrafish GRCz10 genome using TopHat v2.0.13 (ref. 34) with Bowtie v2.2.3.0 (ref. 35), guided by splice junctions defined by cDNA-to-genomic alignments from NCBI Release 104. Transcriptional units were defined separately for each sample using Cufflinks v2.2.1 (ref. 36), and the resulting transcript GTF files were combined between samples (sibling and *clo^{m39/m39}*) and reconciled with NCBI gene annotations using the Cufflinks tool *cuffmerge*. To quantify differential gene expression between sibling and *clo^{m39/m39}* samples, read counts per gene for each sample were calculated using the BEDtools v2.21.0 *intersect* command³⁷ between the alignment BAM file and the exons defined above. Only reads with a MAPQ score > 0 aligning at least 5 nt of an exon were counted. The log₂ fold change between sibling and *clo^{m39/m39}* samples was calculated per gene using counts normalized per million aligned reads (RPM), with a pseudocount of 0.5 added to each count. To generate *P* values, the Cufflinks tool *cuffdiff* was used on the Cufflinks FPKM values of sibling versus *clo^{m39/m39}* samples. Sequencing reads were also aligned to *Danio rerio* EST sequences from GenBank dbEST using Bowtie, and differential expression per EST between sibling and *clo^{m39/m39}* samples was similarly calculated. The genomic loci with reduced expression in *clo^{m39}* mutants are listed in Supplementary Table 2.

Targeted mutagenesis using TALENs and CRISPR/Cas9. TALEN or CRISPR/Cas9 technology was used to target each of the candidate genes identified. TALEN binding sites were selected using TALEN Targeter 2.0 (<https://tale-nt.cac.cornell.edu/>) and constructs were assembled as previously described^{38,39}. RNA was synthesized with mMESSAGE mMachine kit (Ambion) and between 100–150 pg per TALEN arm was injected into one-cell stage embryos to generate potential founders (F₀). For CRISPR/Cas9, sgRNAs for candidate genes were designed using the CRISPR Design online toolbox (<http://crispr.mit.edu>). Guide RNAs were made by ligating annealed primers into the pT7-gRNA vector, and the pT3Ts-nCas9n plasmid (Addgene) was used to synthesize Cas9 mRNA as previously described⁴⁰. To generate founders, between 50–75 pg of sgRNA and 250–500 pg of Cas9 mRNA were coinjected into one-cell stage embryos. To determine the efficacy of individual TALEN or CRISPR/Cas9 injections, DNA was extracted from injected embryos at 48 hpf and subjected to high-resolution melt analysis (HRMA) (Eco-Illumina), and sequencing was used to assess TALEN or CRISPR/Cas9 efficiency. F₀ founders for each gene were identified by outcrossing TALEN or CRISPR/Cas9 injected fish with AB and screening the offspring at 48 hpf using HRMA. PCR products from F₀ founders that showed promising HRMA melt curves were cloned and sequenced, and those that gave interesting mutations (that is, frameshifts) were grown and

crossed to *clo^{m39}* in order to assess the vasculature at 48 hpf. Heterozygous intercrosses of individual *Tg(kdrl:EGFP)* adult F₁s were also used to identify any possible vascular phenotypes as well as obvious morphological defects at 48 hpf. Primer sequences are listed in Supplementary Table 3.

Extension of the *npas4l*-containing scaffold to encompass the whole *npas4l* gene. We combined Sequence Read Archive submission ERA402072 (5 × coverage) with our own PacBio WGS reads (20 × coverage) for a total coverage of 25 × of the zebrafish genome. We used PacBio reads to extend the GRCz10_NA449/KN150101.1 scaffold which we found contains part of the *npas4l* gene. Due to the low coverage and low complexity of the locus, automated assembly using the Celera assembler failed. In a manual approach, all PacBio reads were first mapped to GRCz10 with BWA-MEM⁴¹ using the -x pacbio option. Among the reads mapping to the CRISPR target region, we picked the read that extended farthest beyond the end of the scaffold and manually spliced the extra sequence onto the end of the scaffold. To polish the low-quality extension, we then mapped Illumina WGS reads from Sequence Read Archive submission ERA281351 to GRCz10 with the extended scaffold. We then called variants on the scaffold with FreeBayes (v1.0.0-19-gefe685d)⁴² and ran “bcftools polish” (bcftools v. 1.2) to update the scaffold sequence. Two iterations of this process and manually repairing a single uncorrected deletion in exon 3 yielded the final sequence. Thus, we assembled a genomic sequence to which RNA-seq reads could be mapped to predict exon/intron boundaries, as well as other features of the *npas4l* gene. Predicted splice sites were subsequently tested and confirmed by PCR amplification from AB DNA. We would like to acknowledge the Uppsala Genome Center (NGI, Sweden) for their help with the PacBio sequencing and UPPMAX (SNIC, Sweden) for providing computational resources.

5' and 3' rapid amplification of cDNA ends (RACE). 5' and 3' RACE PCR was performed using the SMARTer RACE cDNA amplification kit (Clontech). Briefly, the following primers were used to amplify the 5' and 3' ends of the *npas4l* cDNA from 24 hpf and tailbud cDNA (Supplementary Table 4); 3'/GSP1 - OAS1323; 3'/GSP2 - OAS1324; 5'/GSP1 - OAS1325; 5'/GSP2 - OAS1326.

Imaging. Zebrafish embryos were immobilized with 0.015% tricaine and mounted in 1% low melting agarose (GeneMate). Brightfield images were acquired using a Nikon SMZ25 and confocal images were acquired using a Zeiss LSM780 with a W Plan-Apochromat 20 × /0.8 DIC objective lens. Images of the heart were acquired using a Zeiss LSM800 Observer with a 20 × /0.8 M27 Plan-Apochromat objective lens.

Whole-mount *in situ* hybridization. Embryos were collected at various developmental stages and fixed with 4% paraformaldehyde in PBS overnight at 4 °C. Whole-mount *in situ* hybridization was performed as described⁴³. Most probes were generated by PCR (primers in Supplementary Table 4). SeqAmp DNA Polymerase (Takara) was used to amplify *npas4l*. T7 was used for transcription and digoxigenin (Roche) for labelling. Embryos were cleared in 80% glycerol and images were acquired using a Nikon SMZ25.

Expression analysis. RT-qPCR was performed in a CFX Connect Real-Time System (BIORAD). RNA was isolated using TRIzol and cDNA was synthesized using the Superscript III First-Strand synthesis system (Invitrogen). Plots show the results of three different experiments (3 pools of embryos per developmental stage or treatment) and each pool consisted of 20–40 embryos. Primers were designed using PerlPrimer⁴⁴, and *gapdh* was used for normalization. Primer sequences are listed in Supplementary Table 4.

Gene expression analysis from published data set. RNA-seq data²⁷ was processed as follows: Time series data available from the Sequence Read Archive under accession SRA048184 was mapped with STAR 2.4.2a⁴⁵ to GRCz10 where scaffold KN150101.1/GRCz10_NA449 was replaced with our extended version. Two replicates are available for each time point. Since plots for both are nearly identical, only the first is shown in Extended Data Fig. 7. Gene-level fragments per kb of exon per million fragments mapped (FPKM) values were computed with Cufflinks 2.2.1 (ref. 46) guided by the Ensembl 83 zebrafish gene annotation, which was manually edited to also include *npas4l*.

Overexpression and rescue studies. Full-length *npas4l* cDNA was amplified from cDNA and cloned into pCS2⁺. mRNA was synthesized using the mMES-SAGE mMACHINE kit and injected at different concentrations into one-cell stage *Tg(kdrl:EGFP)* embryos. Concentrations between 25 and 250 pg were tested to rescue *clo^{s5}* and *clo^{m39}*. As *clo^{s5}* and *clo^{m39}* mutants entirely lack anterior endothelium at 28 hpf, we used the presence of *Tg(kdrl:EGFP)* positive endothelial cells in the anterior region of the embryos to quantify the ability of each mRNA to rescue the *clo* phenotype. For overexpression studies, between 25 and 50 pg of *npas4l* mRNA was injected into AB embryos at the one-cell stage.

Phylogenetic analysis. Diagrams illustrating the domain similarity between murine class I and class II bHLH-PAS proteins and Cloche/Npas4l were assembled on the basis of the bHLH, PAS-A and PAS-B domains of the listed proteins with highly similar results (that is, closest similarity to NPAS4). The online tool

*Phyloeny.fr*⁴⁷ was run with standard settings. The phylogeny of Npas4 and Npas4l was calculated using *Phyloeny.fr* based on the amino acid sequence from the N terminus of the bHLH domain to the C terminus of the PAS-B domain using the listed proteins.

Transactivation domain prediction. The Npas4l protein sequence was analysed for putative transactivation domains using published tools⁴⁸.

Microinjection of morpholinos. The following morpholinos (MOs) were purchased from GeneTools (Philomath, OR) and injected at the indicated amounts: standard control MO (8 ng), CCTCTTACCTCAGTTACAATTTATA; *npas4l* ATG MO (1 ng), GAGTCTCCGAGCTCATCTCACA; *npas4l* I2E3 MO (6 ng), CACCTGGAACACAGTGGAGGATT; *tal1* MO⁴⁹ (6.5 ng), AATGCTCTTACCATCGTTGATTTC; *etv2* MO1 (ref. 50) (4 ng), TTGGTACATTTCCATA TCTTAAAGT; *etv2* MO2 (ref. 50) (4 ng), CACTGAGTCCTTATTTCACTATATC.

Genotyping. High-resolution melt analysis (HRMA) was used to genotype the *npas4l*^{bms59} allele. Embryos or fin-clips were placed in PCR tubes with 50 µl of elution buffer (10 mM Tris-HCl, pH 8.5) containing 1 mg ml⁻¹ proteinase K and incubated at 55 °C for 2 h. The samples were then heated to 95 °C for 10 min to inactivate proteinase K. DyNamo SYBR green (Thermo Fisher Scientific, Waltham, MA) was used to perform PCR and HRMA in an Eco Real-Time PCR System (Illumina, San Diego, CA). PCR protocols: 95 °C for 15 s, then 40 cycles of 95 °C for 2 s, 60 °C for 2 s, and 72 °C for 2 s. Following the PCR, a high-resolution melt curve was generated by collecting SYBR-green fluorescence data in the 65–95 °C range. Genotype analysis was performed on plots of normalized derivative fluorescence. To genotype the *clo*⁵⁵ allele, we performed PCR with primers SR1480 and SR1481 (Supplementary Table 4) using the following conditions: 95 °C for 3 min, then 30 cycles of 95 °C for 30 s, 57 °C for 30 s, and 72 °C for 1 min and a final cycle at 72 °C for 7 min. ExoSAP-IT (Affymetrix) was used to clean up PCR products before Sanger sequencing with primer 1480.

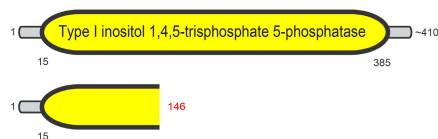
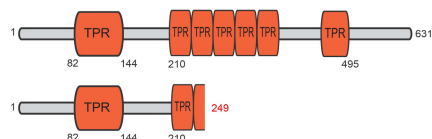
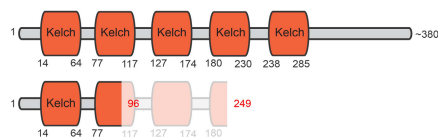
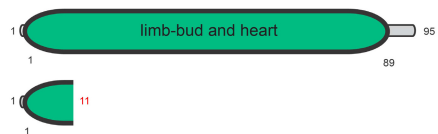
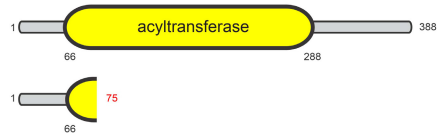
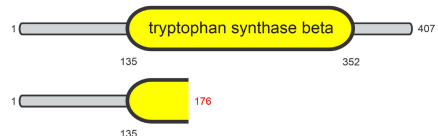
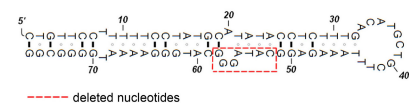
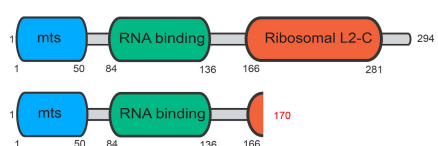
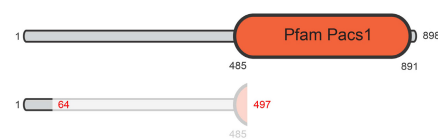
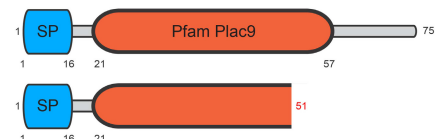
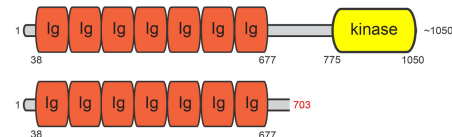
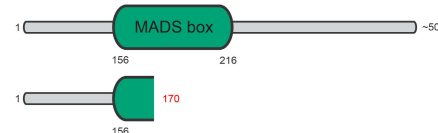
Cell culture. P19 mouse embryonic carcinoma cells (ATCC, CRL-1825) were cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco, Grand Island, NJ) supplemented with 10% fetal bovine serum (FBS, Gibco) and penicillin–streptomycin (Gibco). Mycoplasma tests (Sigma, MP0035) were negative. For transfection, cells were seeded in 6-well plates and transfected with pCS2⁺-AHR or empty pCS2⁺ using FuGene 6. Cells were collected 24 or 48 h after transfection. RNA was isolated using TRIzol and cDNA was synthesized using the Superscript III First-Strand synthesis system.

Endothelial differentiation of human embryonic stem cells. The human embryonic stem cell line H9-hTnTZ-pGZ-D2 (WiCell) was maintained as previously described⁵¹. Briefly, human embryonic stem cells were maintained on Geltrex (ThermoFisher Scientific)-coated plates in E8 media. Human embryonic stem cells were differentiated towards the vascular endothelial lineage as previously described^{52,53}. On day 0 of differentiation, human embryonic stem cells were dissociated using Versene (ThermoFisher Scientific) and embryoid bodies were generated in StemPro-34 media (ThermoFisher Scientific) in the presence of recombinant human (rh) BMP4 (1 ng ml⁻¹, ThermoFisher Scientific). After 24 h, embryoid bodies were cultured in StemPro-34 supplemented with rhBMP4 (6 ng ml⁻¹), rhACTIVINA (4 ng ml⁻¹, ThermoFisher Scientific), and rhbFGF (5 ng ml⁻¹, PeproTech) to induce mesoderm formation. On day 4 of differentiation, embryoid bodies were dissociated using TrypLE (Invitrogen) and plated in monolayer on Geltrex-coated plates at 1 × 10⁵ cells per well in StemPro-34 supplemented with rhVEGF (100 ng ml⁻¹) and rhbFGF (50 ng ml⁻¹). Cultures were maintained in this media until day 10 of differentiation with media changes every 2 days.

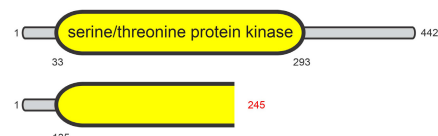
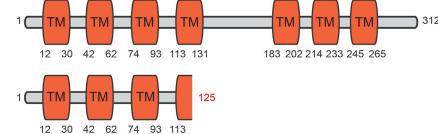
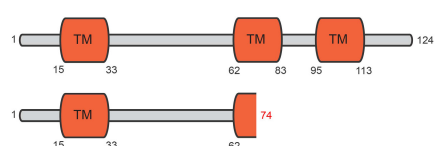
Cell staining for flow cytometry was carried out in PBS with 0.5% BSA. Cells were stained at a concentration of 2.5 × 10⁶ cells per ml with anti-KDR-PE (R&D Systems; 1:10), anti-PDGFRα-APC (R&D Systems; 1:10), and anti-CD31-PE-Cy7 (BioLegend; 1:100). Stained cells were analysed on a Guava 8HT flow cytometer (EMD Millipore).

For RT-qPCR, total RNA was prepared with the Direct-Zol MiniPrep kit (Zymo Research). RNA was reverse transcribed into cDNA using random hexamers and Oligo(dT) with Superscript III Reverse Transcriptase (Life Technologies, Invitrogen). RT-qPCR was performed on a Bio-Rad CFX Connect Real-Time PCR Detection System (Bio-Rad) using Power SYBR Green Master Mix (Life Technologies, Applied Biosystems). Expression levels were normalized to the housekeeping gene TATA box binding protein (*TBP*). Primer sequences are listed in Supplementary Table 4.

- Jin, S. W., Beis, D., Mitchell, T., Chen, J. N. & Stainier, D. Y. Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. *Development* **132**, 5199–5209 (2005).
- Yi, H. et al. Duplex-specific nuclease efficiently removes rRNA for prokaryotic RNA-seq. *Nucleic Acids Res.* **39**, e140 (2011).
- Zhao, W. et al. Comparison of RNA-seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genomics* **15**, 419 (2014).
- Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
- Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
- Roberts, A., Pimentel, H., Trapnell, C. & Pachter, L. Identification of novel transcripts in annotated genomes using RNA-seq. *Bioinformatics* **27**, 2325–2329 (2011).
- Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
- Cermak, T. et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39**, e82 (2011).
- Bedell, V. M. et al. In vivo genome editing using a high-efficiency TALEN system. *Nature* **491**, 114–118 (2012).
- Jao, L. E., Wente, S. R. & Chen, W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc. Natl Acad. Sci. USA* **110**, 13904–13909 (2013).
- Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at <http://arxiv.org/abs/1303.3997v2> (2013).
- Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. Preprint at <http://arxiv.org/abs/1207.3907> (2012).
- Thisse, C. & Thisse, B. High-resolution *in situ* hybridization to whole-mount zebrafish embryos. *Nature Protocols* **3**, 59–69 (2008).
- Marshall, O. J. PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. *Bioinformatics* **20**, 2471–2472 (2004).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
- Trapnell, C. et al. Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
- Dereeper, A. et al. Phyloeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36**, W465–W469 (2008).
- Piskacek, S. et al. Nine-amino-acid transactivation domain: establishment and prediction utilities. *Genomics* **89**, 756–768 (2007).
- Dooley, K. A., Davidson, A. J. & Zon, L. I. Zebrafish *scl* functions independently in hematopoietic and endothelial development. *Dev. Biol.* **277**, 522–536 (2005).
- Sumanas, S. & Lin, S. Ets1-related protein is a key regulator of vasculogenesis in zebrafish. *PLoS Biol.* **4**, e10 (2006).
- Chen, G. et al. Chemically defined conditions for human iPSC derivation and culture. *Nature Methods* **8**, 424–429 (2011).
- Kattman, S. J. et al. Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell Stem Cell* **8**, 228–240 (2011).
- Lin, B. et al. High-purity enrichment of functional cardiovascular cells from human iPS cells. *Cardiovasc. Res.* **95**, 327–335 (2012).

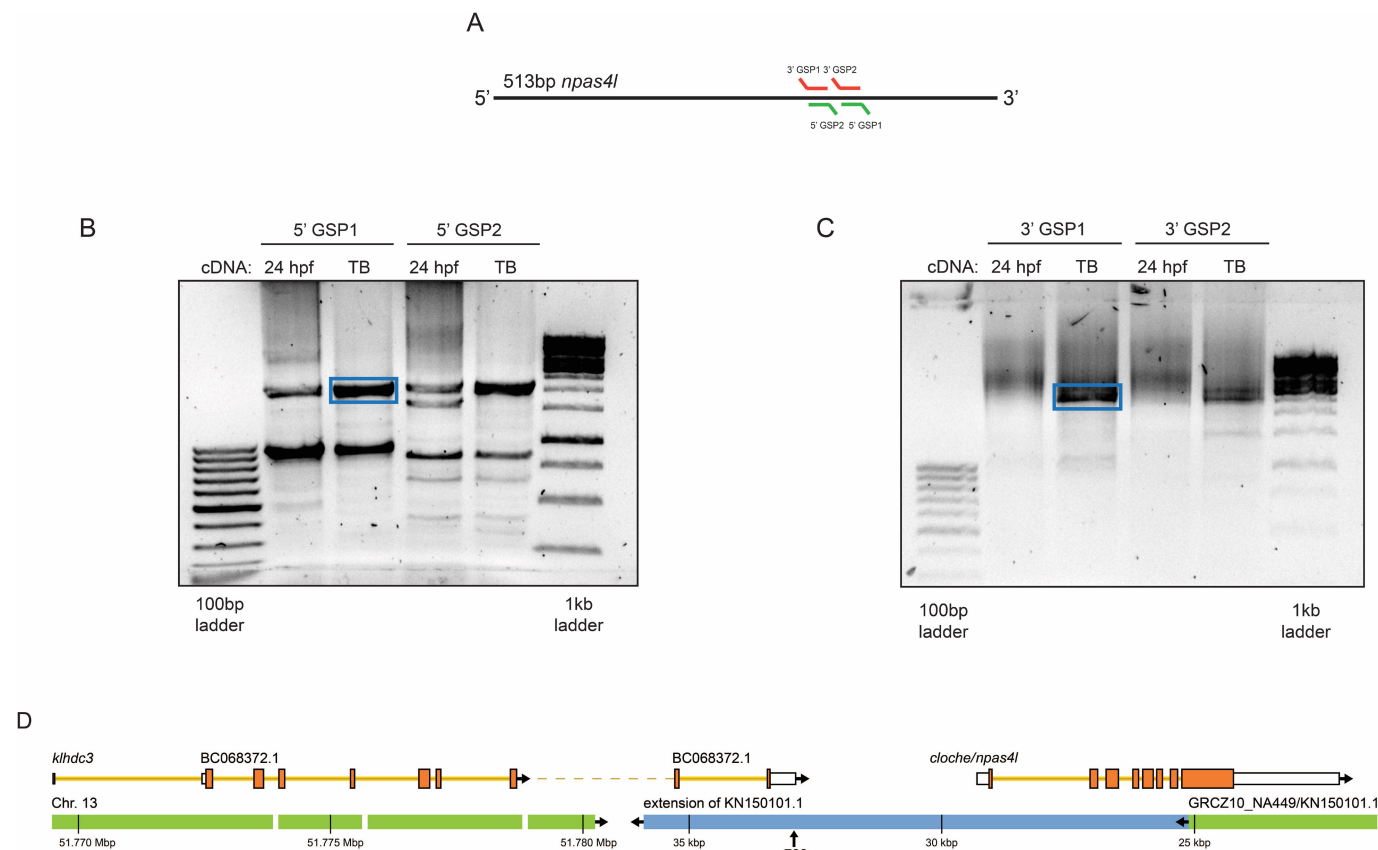
Inpp5a - $\Delta 14$ bpKlc1b - $\Delta 10$ bpKlhdc3 - $\Delta 5$ bp; $\Delta 1$ bpLbh - $\Delta 4$ bpLclat1 - $\Delta 4$ bpLOC100331665 - $\Delta 25$ bpmir202 - $\Delta 7$ bpMrpl2 - $\Delta 20$ bpPacs1b - $\Delta 13$ bp; $\Delta 17$ bpPlac9 - $\Delta 8$ bpPtk7b - $\Delta 7$ bpSlc30a10 - $\Delta 22$ bpSrff - $\Delta 8$ bp

Stk32c - +8bp

Tmem121-like - $\Delta 24$ bpTmem254 - $\Delta 4$ bp

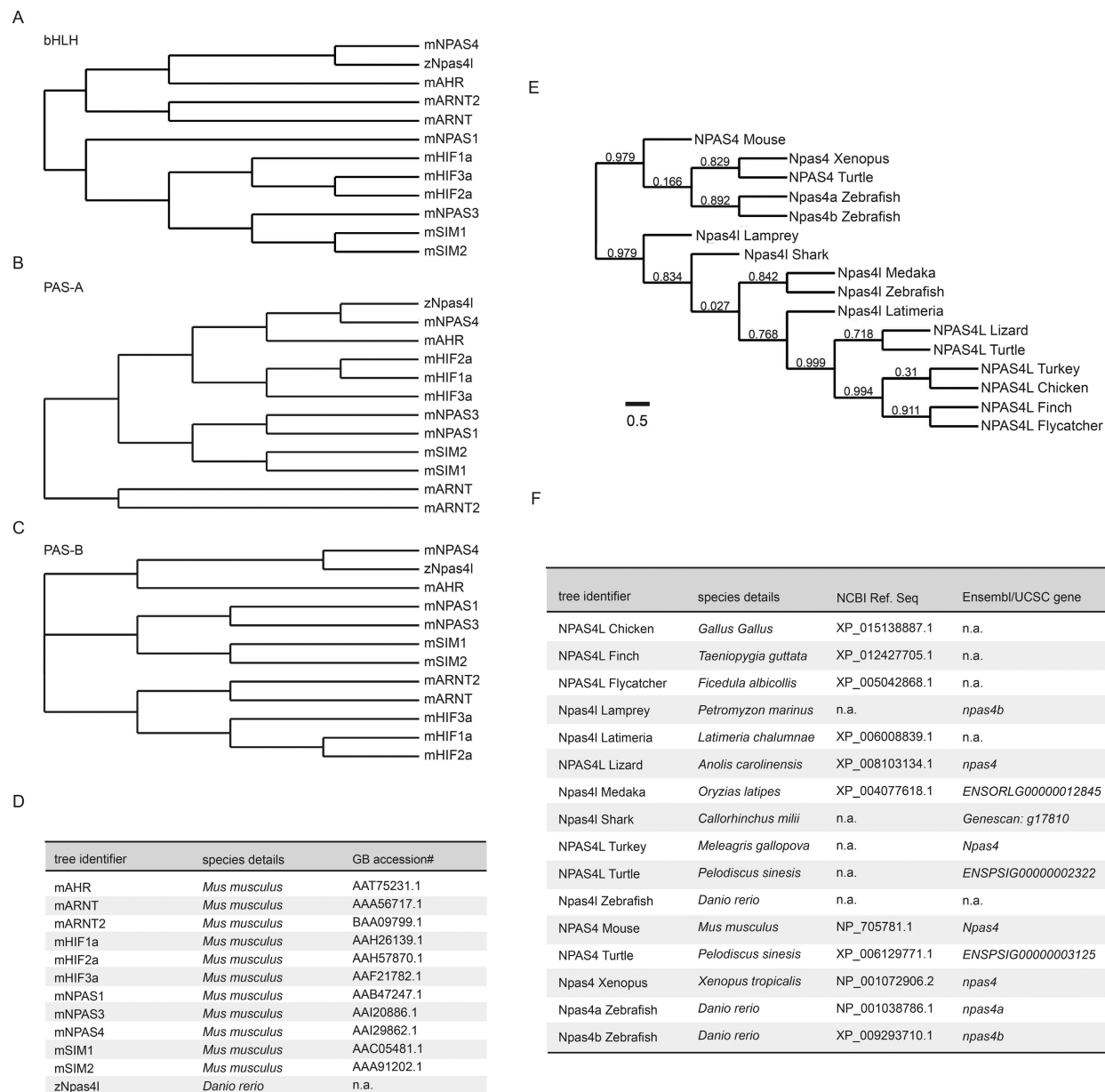
Extended Data Figure 1 | Illustration of the predicted mutant proteins. Schematic representations show the wild-type protein (top) and the corresponding mutant protein (bottom) for each of the candidate genes in the *cloche* region. Functional protein domains are annotated, and the information about the molecular lesion(s) is indicated next to the protein's name. Red numbers indicate the last native amino acid before the induced frameshift mutation. For the two genes that were targeted

at more than one position, the extension of the longer mutant protein is indicated by transparency. Except for the *npas4l* mutation, none of the other 18 mutations causes an obvious morphological phenotype at 48 hpf. For proteins with incomplete sequence information (Inpp5a, Ptk7b, Srff), primary structure was extrapolated from orthologues from other teleost species.



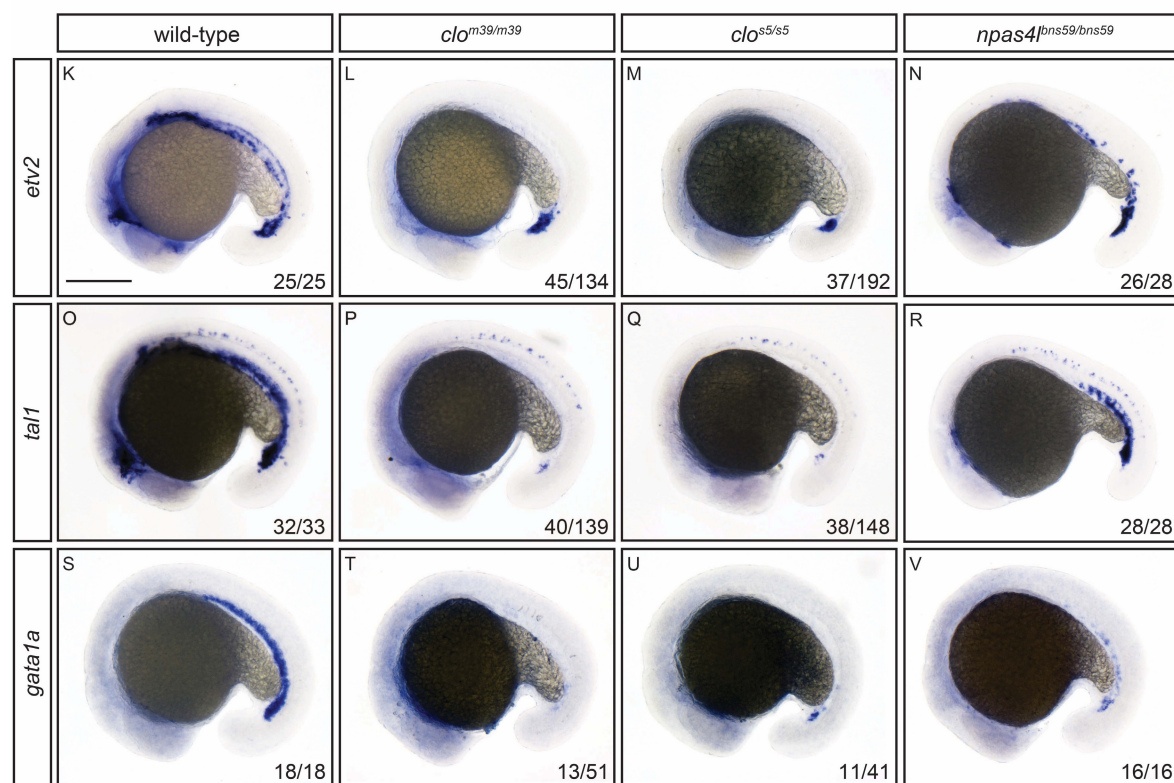
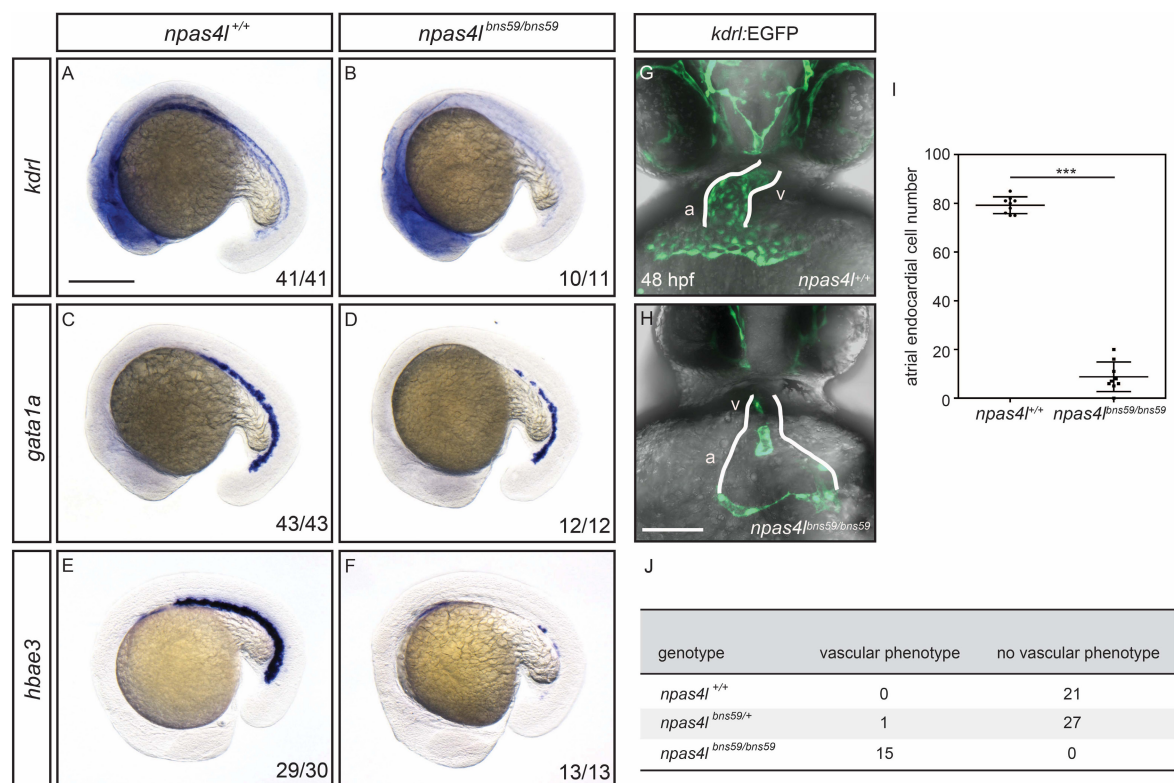
Extended Data Figure 2 | 5' and 3' RACE PCR for *npas4l*, and physical map of the *cloche/npas4l* locus in relation to *klhdc3*. **a**, Primer design was based on a 513 bp fragment from GRCZ10 that is deleted in *clo*^{m39} mutants. **b**, **c**, Agarose gels showing the products of 5' (**b**) and 3' (**c**) RACE PCRs using cDNAs generated from 24 hpf or tailbud (TB) mRNAs. Multiple fragments were cloned for sequencing and bands that extended the *npas4l* sequence to the 5' (**b**) and 3' (**c**) are highlighted in blue boxes. **d**, Physical

assembly of the *cloche* locus based on the extension of KN150101.1 by whole-genome sequencing. KN150101.1 (ex) is shown in reverse-complementary orientation with the blue region marking newly obtained sequence information. Alignment of the *klhdc3* expressed sequence tag (EST) BC068372.1 establishes a physical connection between KN150101.1 (ex) and Chr. 13. The tightly linked SSLP marker z709 lies between *klhdc3* and *cloche/npas4l*.



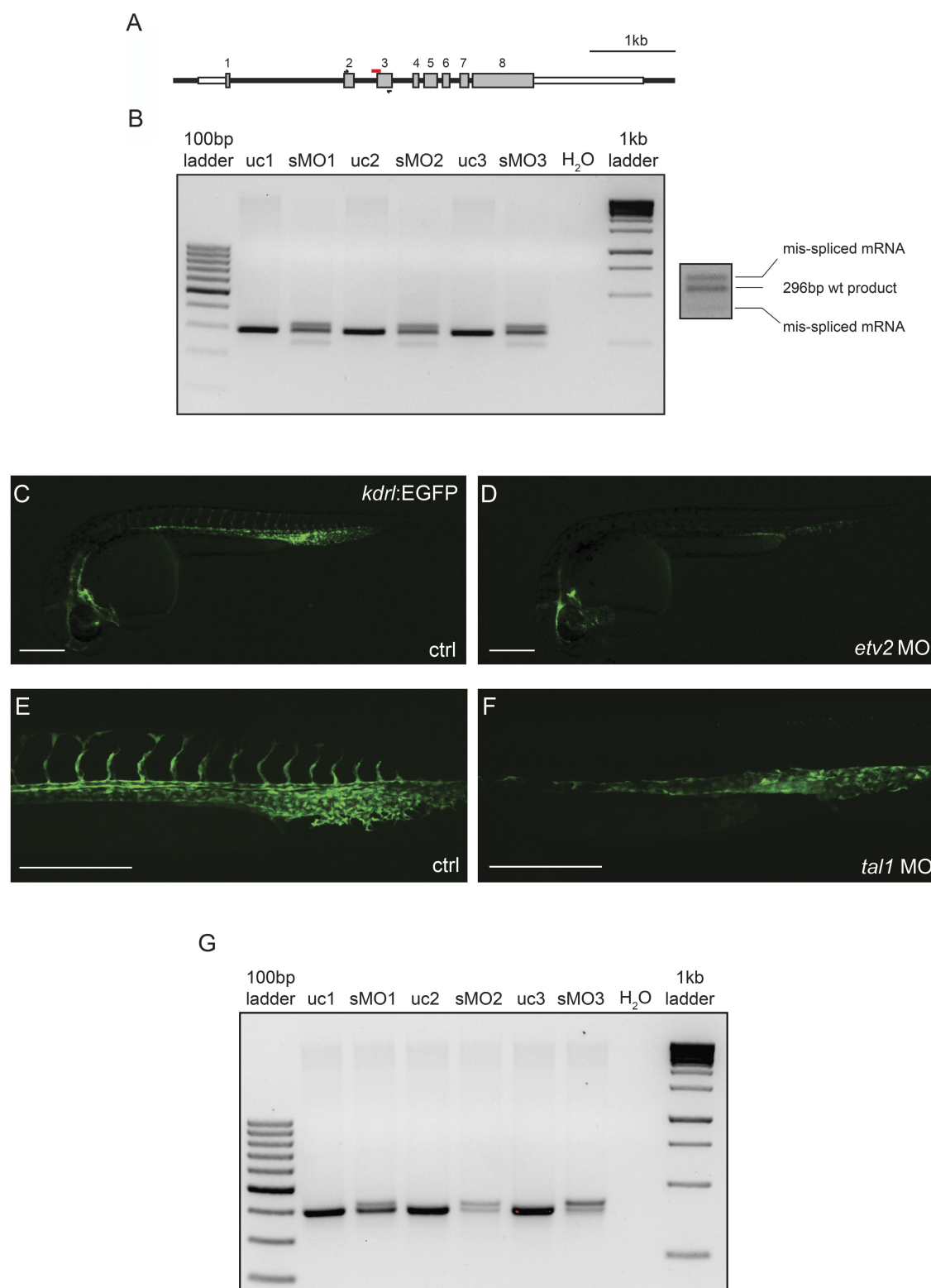
Extended Data Figure 3 | bHLH-PAS protein relatedness and phylogeny of Cloche/Npas4l. **a–c**, Individual clustering diagrams based on the amino acid sequence of the bHLH domain (**a**), the PAS-A domain (**b**), and the PAS-B domain (**c**) of mouse class I (NPAS1, NPAS3, NPAS4, HIF1A-HIF3A, AHR, SIM1/2) and class II (ARNT, ARNT2) bHLH-PAS proteins, together with the protein encoded by zebrafish *cloche/npas4l*. Analysis using the respective zebrafish proteins leads to similar results; the similarity between full-length zebrafish Npas4l and Npas4a is 39.5%, and between Npas4l and Npas4b it is 33%. **d**, GenBank accession numbers

of proteins used in this analysis. **e**, Unrooted phylogenetic tree of Npas4 and Npas4l proteins across different phyla (using the sequence from the N terminus of the bHLH domain to the C terminus of the PAS-B domain). The *cloche/npas4l* gene is clearly present from lampreys to birds, but appears to be missing in mammals. **f**, Detailed information on the NPAS4 and NPAS4L protein sequences as well as synteny information of the corresponding genes. Predicted proteins were manually corrected when necessary using gene prediction methods (Genscan; AUGUSTUS) and BLAST tools.



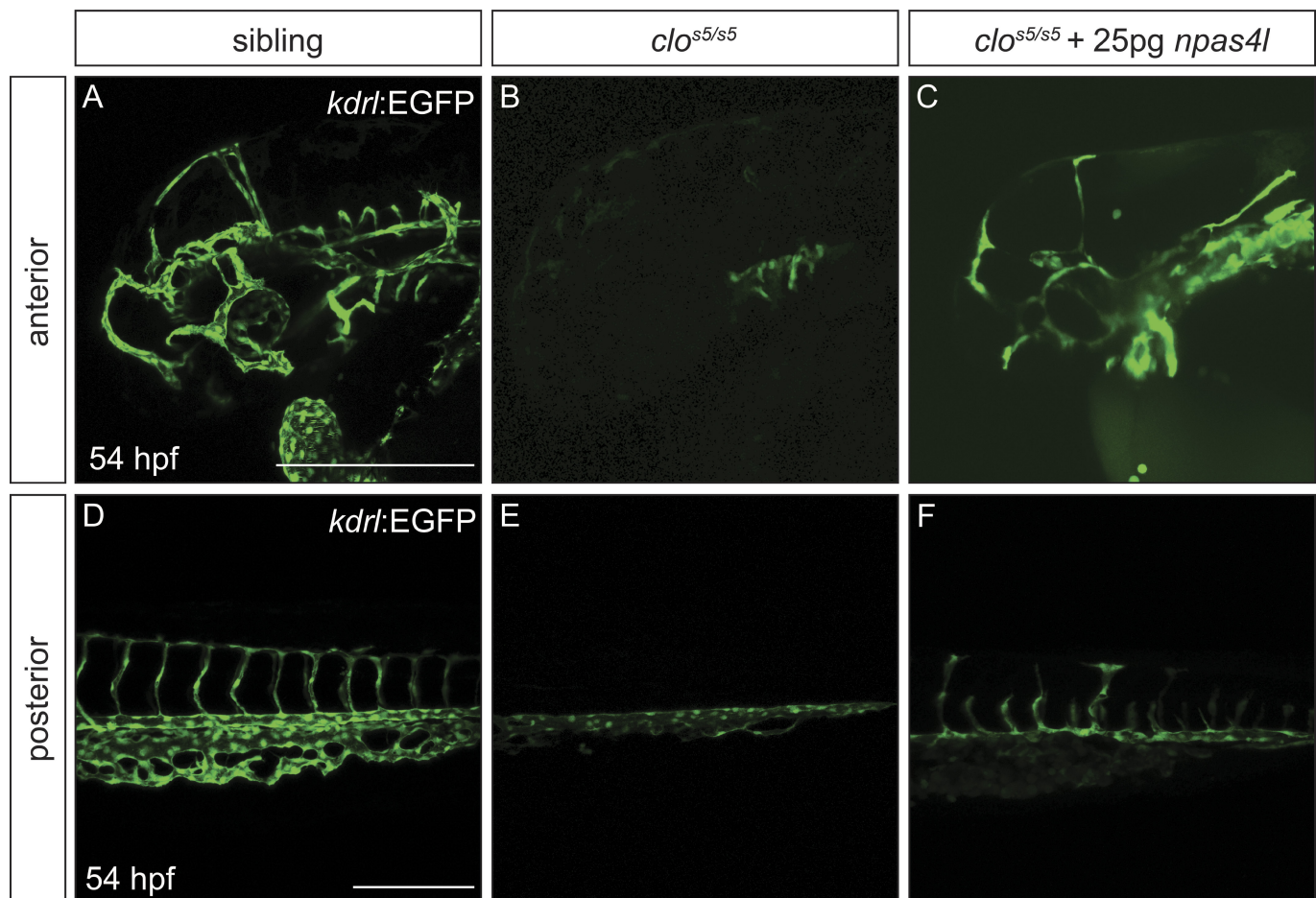
Extended Data Figure 4 | Additional phenotypic analysis of *npas4l*^{bns59} mutant. **a–f**, Whole-mount *in situ* hybridization for *kdr1* (**a**, **b**), *gata1a* (**c**, **d**) and *hbae3* (**e**, **f**) expression at the 15-somite stage in wild-type and *npas4l*^{bns59/bns59} embryos. **g**, **h**, Ventral views of *Tg(kdr1:EGFP)* sibling (**g**) and *npas4l*^{bns59} mutant (**h**) hearts at 48-hpf; a, atrium; v, ventricle. **i**, Quantification of the number of atrial endocardial cells in homozygous wild-type siblings ($n = 9$) and *npas4l*^{bns59} mutants ($n = 9$) at 60 hpf. Values represent means \pm s.e.m. *** $P < 0.001$ by *t*-test. **j**, Penetrance of the vascular phenotype based on *Tg(kdr1:EGFP)* expression in 48 hpf

npas4l^{bns59/bns59} embryos. **k–v**, Side-by-side comparison of endothelial and haematopoietic marker gene expression between different *cloche* alleles. Whole-mount *in situ* hybridization for *etv2* (**k–n**), *tal1* (**o–r**) and *gata1a* (**s–v**) expression at the 15-somite stage in wild-types and *clo*^{m39}, *clo*^{s5} and *npas4l*^{bns59} mutants. *x/y*, number of embryos showing representative phenotype (*x*), number of embryos examined (*y*); *m39* and *s5* mutants were not confirmed by genotyping, hence the approximately 1/4 ratio in those panels. Scale bars, 200 μ m (except for **g** and **h**, 100 μ m).

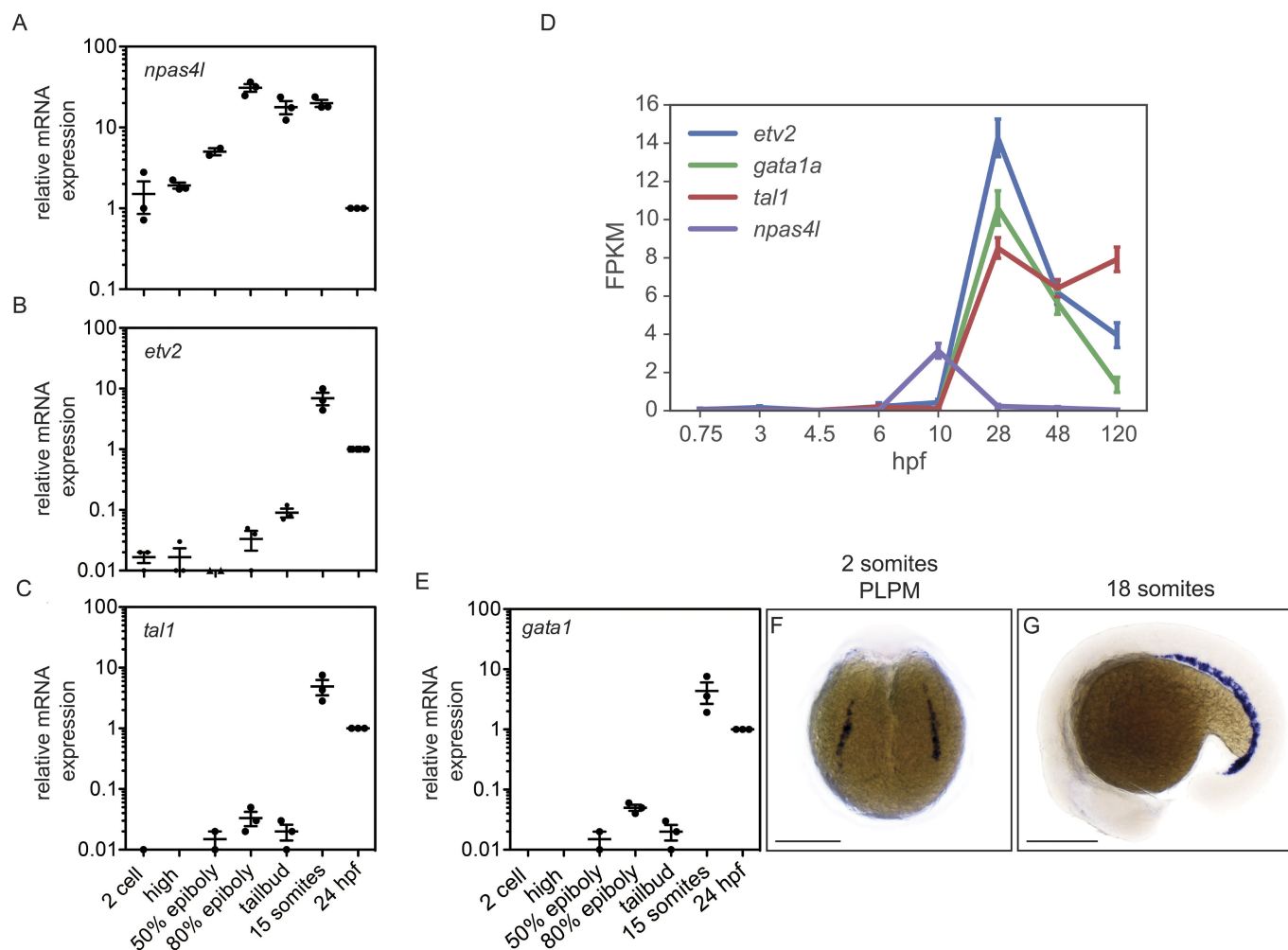


Extended Data Figure 5 | Morpholino knockdown of *npas4l*, *etv2* and *tal1*. Knockdown of *Npas4l* was achieved by co-injecting a translation blocking morpholino targeting the *npas4l* translational start site and a splice morpholino targeting the intron-exon splice site I2E3. **a**, Schematic representation of the *npas4l* genomic locus including the I2E3 splice morpholino (red) and the oligos used to test for splicing defects by RT-PCR (black). **b**, Analysis of *npas4l* mRNA splicing by RT-PCR shows a significant reduction in wild-type *npas4l* mRNA at the tailbud stage following injection of 6 ng of the I2E3 morpholino at the one-cell stage (3 different experiments, sMO1–sMO3, are shown). uc, uninjected control; sMO, splice morpholino injected; H₂O, water control; wild-type

product, 296 bp. **c**, **d**, Knockdown of *Etv2* was achieved by injecting two previously published ATG morpholinos⁵⁰. Fluorescence micrographs of *Tg(kdrl:EGFP)* control (ctrl) and *etv2* MO1 + MO2 injected embryos. **d–f**, Knockdown of *Tal1* was achieved by injecting a previously published splice morpholino⁴⁹. **d**, **e**, Confocal projections of *Tg(kdrl:EGFP)* ctrl and *tal1* MO injected embryos. **g**, Analysis of *tal1* mRNA splicing by RT-PCR shows a significant reduction in wild-type *tal1* mRNA at the two-somite stage after injecting 6.5 ng of *tal1* MO at the one-cell stage. Three different experiments, sMO1–sMO3, are shown). uc, uninjected control; sMO, splice morpholino injected; H₂O, water control; wild-type product, 401 bp. Scale bars, 200 μ m.

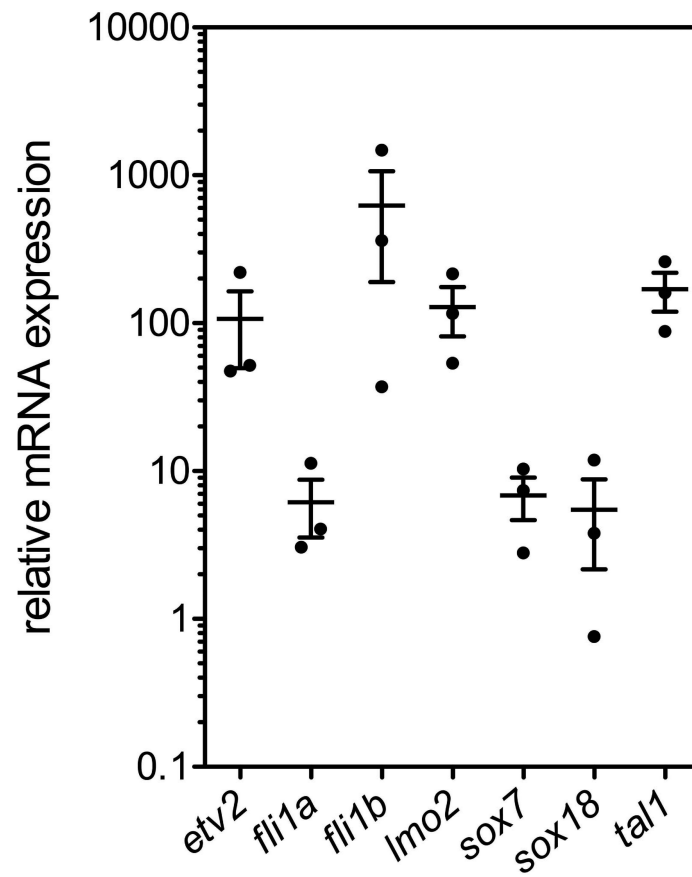


Extended Data Figure 6 | *npas4l* rescues *clo*^{s5/s5} embryos. a–f, Confocal projections of the anterior (a–c) and posterior (d–f) vasculature of 54 hpf *Tg(kdr*l:EGFP) sibling, *clo*^{s5/s5} and *clo*^{s5/s5} embryos injected with 25 pg *npas4l* mRNA at the one-cell stage. Scale bars, 200 μ m.

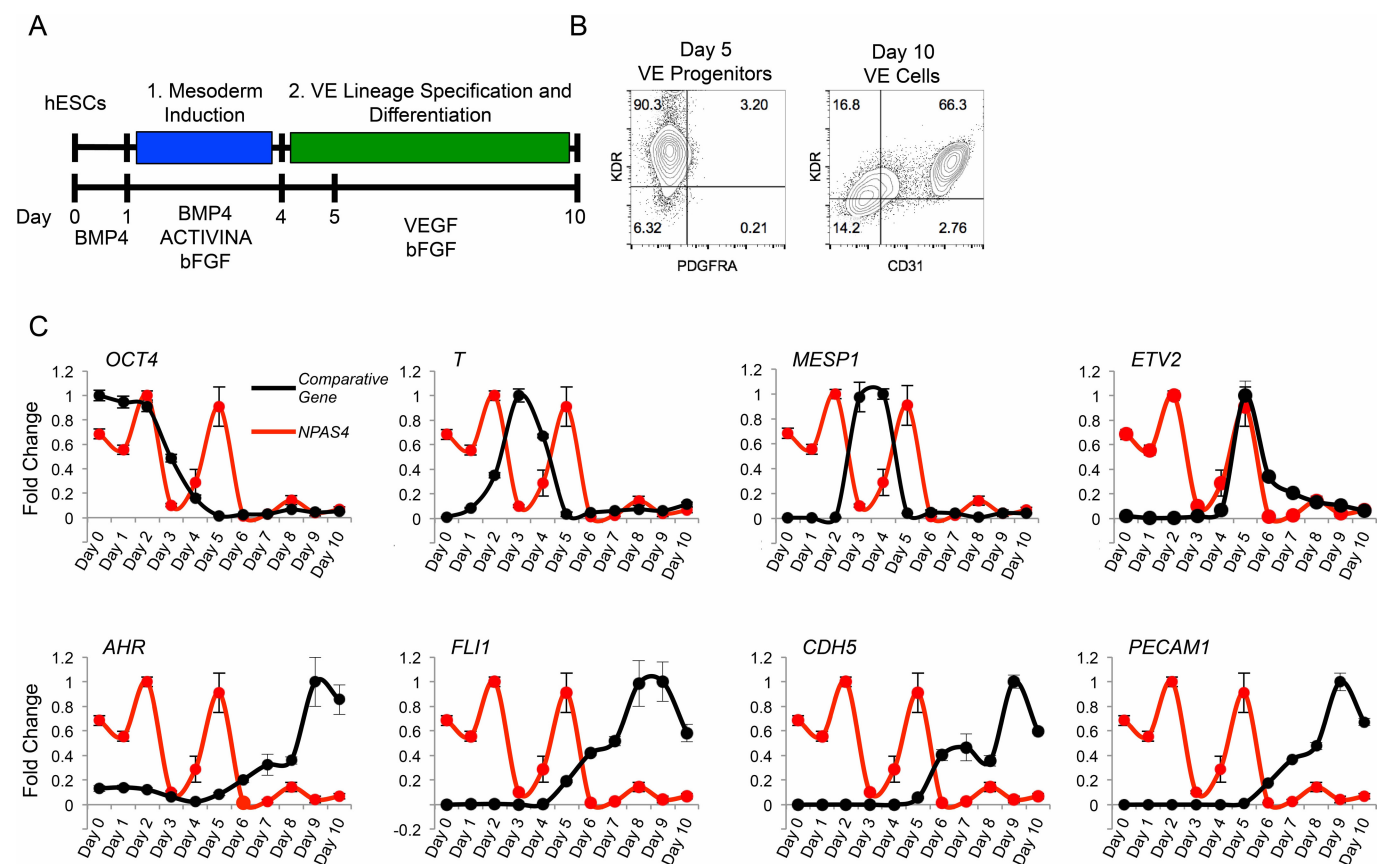


Extended Data Figure 7 | Expression levels of *npas4l*, *etv2*, *tal1* and *gata1a* during development. **a–c**, RT–qPCR for *npas4l*, *etv2* and *tal1* expression; 20–40 embryos from single pair matings were pooled at each of the indicated stages; expression is normalized to the corresponding 24 hpf sample. Error bars represent s.e.m. from three biological replicates. **d**, Data extracted from ref. 27. *npas4l* expression clearly precedes *etv2*, *tal1* and *gata1a* expression; it appears to be minimal until 6 hpf, peaks around the end of gastrulation/early somitogenesis (in 10 hpf

sample from this data set), and is clearly down by 28 hpf, the stage when the other transcripts peak. Expression at each developmental stage represents a single RNA-seq experiment on 1,000 pooled embryos, and bars indicate the 95% confidence interval (Cufflinks). **e**, RT–qPCR for *gata1a* expression, performed and analysed as in **a–c**. Whole-mount *in situ* hybridization for *gata1a* expression in wild-type embryos at the 2-somite (**f**) and 18-somite (**g**) stages. Scale bars, 200 μ m.

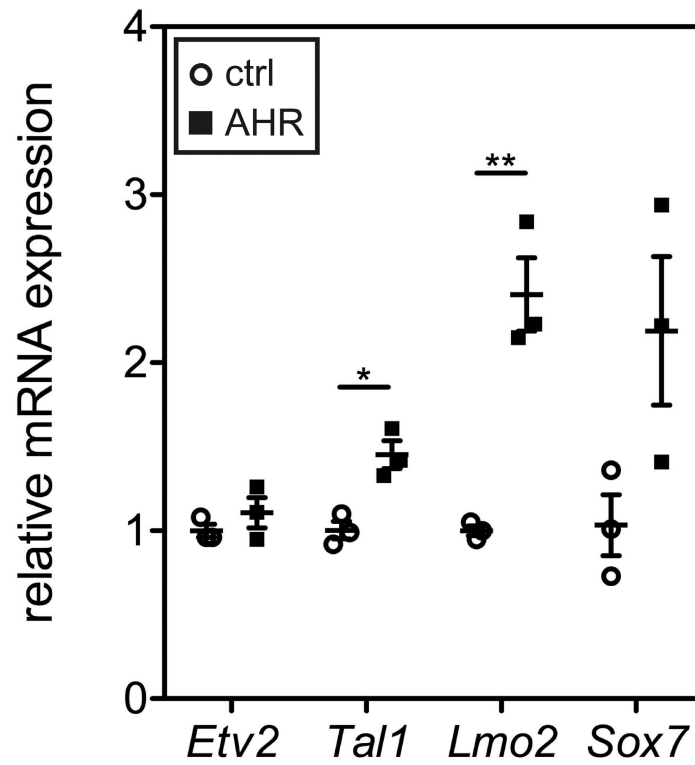


Extended Data Figure 8 | Potential *Npas4l* targets. RT-qPCR analysis of *etv2*, *fli1a*, *fli1b*, *lmo2*, *sox7*, *sox18* and *tal1* expression in shield stage embryos injected with *npas4l* mRNA at the one-cell stage, normalized to uninjected controls. Error bars represent s.e.m. from three biological replicates of 20–40 pooled embryos per condition.



Extended Data Figure 9 | *NPAS4* expression during human embryonic stem cell derived vascular endothelial differentiation. **a**, Scheme of the protocol used to differentiate human embryonic stem cells towards the vascular endothelial (VE) lineage highlighting the (1) mesoderm induction stage and (2) the VE lineage specification and differentiation stage. **b**, Representative flow cytometric analysis for day 5 KDR^{+} $PDGFR\alpha^{-}$ VE progenitors and day 10 $CD31^{+}$ KDR^{+} VE cells.

c, RT-qPCR analysis of temporal expression of *NPAS4* compared to *OCT4*, *T*, *MESP1*, *ETV2*, *AHR*, *FLI1*, *CDH5* and *PECAM1* over the course of human embryonic stem cell derived VE differentiation. Values are relative to the housekeeping gene *TBP* and depicted as a fold change in which average peak expression of the individual gene is set to 1. Raw data can be found in Figshare (<http://dx.doi.org/10.6084/m9.figshare.3383572>). Error bars represent s.e.m. from three biological replicates.



Extended Data Figure 10 | AHR overexpression induces expression of early endothelial and haematopoietic markers in P19 mouse embryonic carcinoma cells. a, RT-qPCR analysis of *Etv2*, *Tal1*, *Lmo2* and *Sox7* expression following overexpression of human AHR in P19 cells. Error bars represent s.e.m. from three biological replicates derived from independent transfections (* $P < 0.05$ and ** $P < 0.01$ by *t*-test).

Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios

Philipp S. Hoppe^{1,2}, Michael Schwarzfischer³, Dirk Loeffler^{1,2}, Konstantinos D. Kokkaliaris^{1,2}, Oliver Hilsenbeck^{1,2,3}, Nadine Moritz², Max Endeke^{1,2}, Adam Filipczyk², Adriana Gambardella⁴, Nouraiz Ahmed¹, Martin Etzrodt¹, Daniel L. Coutu¹, Michael A. Rieger², Carsten Marr³, Michael K. Strasser³, Bernhard Schaubberger², Ingo Bartscher⁵, Olga Ermakova⁶, Antje Bürger⁷, Heiko Lickert^{5,8}, Claus Nerlov^{4,9}, Fabian J. Theis^{3,10} & Timm Schroeder^{1,2}

The mechanisms underlying haematopoietic lineage decisions remain disputed. Lineage-affiliated transcription factors^{1,2} with the capacity for lineage reprogramming³, positive auto-regulation^{4,5} and mutual inhibition^{6,7} have been described as being expressed in uncommitted cell populations⁸. This led to the assumption that lineage choice is cell-intrinsically initiated and determined by stochastic switches of randomly fluctuating cross-antagonistic transcription factors³. However, this hypothesis was developed on the basis of RNA expression data from snapshot and/or population-averaged analyses^{9–12}. Alternative models of lineage choice therefore cannot be excluded. Here we use novel reporter mouse lines and live imaging for continuous single-cell long-term quantification of the transcription factors GATA1 and PU.1 (also known as SPI1). We analyse individual haematopoietic stem cells throughout differentiation into megakaryocytic–erythroid and granulocytic–monocytic lineages. The observed expression dynamics are incompatible with the assumption that stochastic switching between PU.1 and GATA1 precedes and initiates megakaryocytic–erythroid versus granulocytic–monocytic lineage decision-making. Rather, our findings suggest that these transcription factors are only executing and reinforcing lineage choice once made. These results challenge the current prevailing model of early myeloid lineage choice.

Multipotent haematopoietic stem and progenitor cells (HSPCs) are thought to differentiate into all blood cell types through a series of progenitor cell types with increasingly restricted lineage potential—for example, the common myeloid progenitor (CMP), which then further differentiates into megakaryocytic–erythroid (MegE) and granulocytic–monocytic (GM) progenitors (MEPs and GMPs)¹³. The molecular mechanisms controlling lineage choice remain controversial. The prevailing model assumes that lineage choice is initiated and determined by stochastic fluctuations of cross-antagonistic transcription factor (TF) pairs^{3,14}, developed around the haematopoietic TFs PU.1 and GATA1. PU.1 and GATA1 are expressed in GM and MegE cells¹³, respectively; they are required for the production of mature cells of these lineages^{1,2} and can reprogram cells towards their lineages upon overexpression³. PU.1 and GATA1 can cross-inhibit each other's activity^{6,7} and activate their own transcription^{4,5}. This wiring can generate bi-stable switches, with random higher expression of one TF, leading to increased expression of this TF and repression of competing TFs. *PU.1* and *Gata1* mRNAs were described as being co-expressed before HSPC lineage choice⁸. Therefore, lineage decisions may be initiated by random fluctuations of TF levels, breaking the TF equilibrium of a cell. The specific wiring of the TF network would lead to specific probabilities

for individual TFs to 'win', and thus to stable frequencies of lineage choices. In this model, TFs would not only execute and reinforce, but also initiate and make lineage decisions¹⁴.

However, this model was based on low-resolution expression data that cannot be used to exclude alternative models. Most studies only analysed HSPC population averages^{9–11}, thus masking cellular heterogeneity^{12,15} and dead cells to non-quantitatively measure mRNA expression, ignoring relevant protein expression dynamics and future lineage choice^{12,16}. We therefore developed approaches for continuous long-term single-cell quantification of PU.1 and GATA1 protein expression in individual HSPCs from before until after their lineage choice. We created knock-in mouse lines with reading frames for yellow (enhanced yellow fluorescent protein; eYFP) and red (mCherry) fluorescent proteins knocked into the gene loci for *PU.1* and *Gata1*, respectively (Fig. 1a, Extended Data Fig. 1a). The fluorescent proteins are fused to the C terminus of each TF, allowing quantification by fluorescence intensity. We mated the resulting *PU.1*^{eYFP} (ref. 17) and *GATA1*^{mCherry} mice to create *PU.1*^{eYFP}*GATA1*^{mCherry} mice. These are homozygous for both *PU.1*^{eYFP} and *Gata1*^{mCherry} alleles, and males are hemizygous for X-chromosomal *Gata1*^{mCherry}.

As described previously, no² or reduced¹⁸ PU.1 expression or altered PU.1 function¹⁹, and no¹ or altered GATA1 expression²⁰ have drastic phenotypes, causing leukaemia or absence of mature GM or MegE cells, for example. In contrast, *PU.1*^{eYFP}*GATA1*^{mCherry} mice show no aberrant phenotypes, are born at normal Mendelian ratios (Extended Data Table 1), and did not show increased mortality throughout adulthood (data not shown). The cellular composition of GATA1–mCherry-expressing fetal livers (Extended Data Fig. 1b), peripheral blood (Extended Data Fig. 1c) and bone marrow (Fig. 1b–f) from adult *PU.1*^{eYFP}*GATA1*^{mCherry} mice was unchanged. Colony formation *in vitro* was unaltered for *PU.1*^{eYFP}*GATA1*^{mCherry} cells (Fig. 1g, h). We could also not observe a difference in GM to MegE lineage output of *PU.1*^{eYFP}*GATA1*^{mCherry} versus wild-type HSPCs in competitive repopulation assays (Fig. 1i, j, Extended Data Fig. 2). Finally, reprogramming of cells to MegE or GM lineages by GATA1–mCherry or PU.1–eYFP, respectively, was as efficient as with wild-type TFs, both in wild-type and in *PU.1*^{eYFP}*GATA1*^{mCherry} cells (Extended Data Fig. 3). In conclusion, PU.1–eYFP and GATA1–mCherry fusion proteins function normally.

We showed normal expression and stability of the fusion proteins by quantitative immunofluorescence staining for PU.1 and GATA1. Simultaneous staining for the TFs and the fluorescent proteins (Fig. 1k–n) showed high expression correlation and co-localization in HSPC nuclei. Distributions of PU.1 or GATA1 protein expression were

¹Department of Biosystems Science and Engineering, ETH Zurich, 4058 Basel, Switzerland. ²Research Unit Stem Cell Dynamics, Helmholtz Zentrum München, 85764 Neuherberg, Germany.

³Institute of Computational Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany. ⁴MRC Molecular Haematology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford OX3 9DS, UK. ⁵Institute of Diabetes and Regeneration Research, Institute of Stem Cell Research, Helmholtz Zentrum München, Business Campus Garching, 85748 Garching, Germany. ⁶Istituto di Biologia Cellulare e Neurobiologia, CNR, 00015 Monterotondo, Italy. ⁷Institute of Developmental Genetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany.

⁸Beta Cell Biology, Medical Faculty, Technical University Munich (TUM), 80333 Munich, Germany. ⁹EMBL Mouse Biology Unit, 00015 Monterotondo, Italy. ¹⁰Department of Mathematics, Technical University Munich, 85748 Garching, Germany.

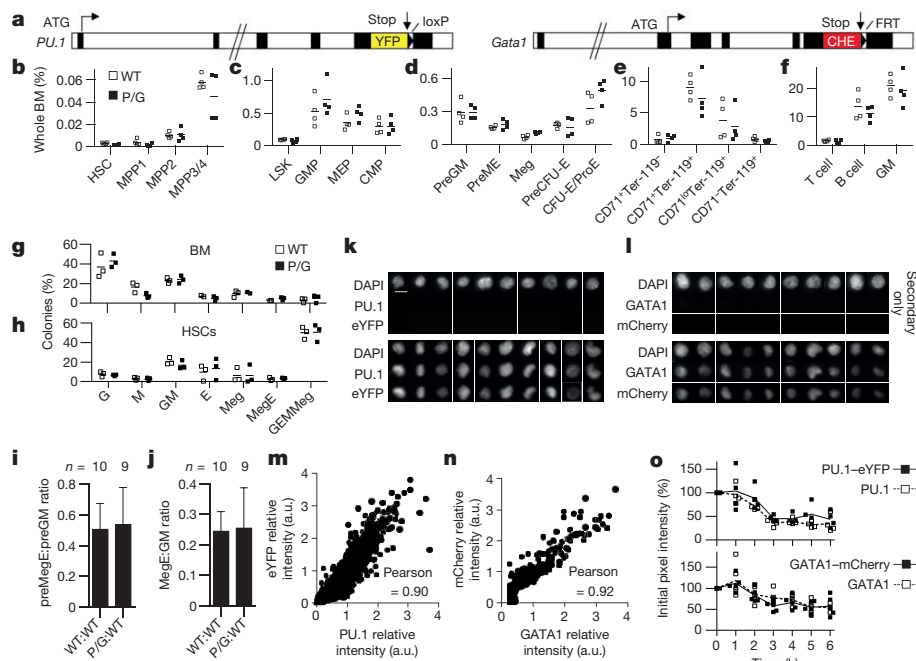


Figure 1 | Normal function and stability of PU.1-eYFP and GATA1-mCherry. **a**, Endogenous gene loci after knock-in. Black boxes, exons. **b-f**, Normal haematopoiesis in PU.1^{eYFP}GATA1^{mCherry} (P/G) double homozygous mice. Composition of adult bone marrow (4 independent experiments). One-way MANOVA for **b-f**, $P = 0.35, 0.35, <0.03, 0.31, 0.16$, respectively. **E**, erythroid. **g-j**, Normal lineage choice of PU.1^{eYFP}GATA1^{mCherry} HSPCs. **g, h**, Colony-forming assay from whole bone marrow (BM) (**g**) and HSCs (**h**) (mean, 3 independent experiments each). No significant difference for each population (Wilcoxon rank sum test, P values >0.2 and >0.4 , respectively).

not changed in PU.1^{eYFP}GATA1^{mCherry} HSPC populations (Extended Data Figs 4, 5a–e), and resembled those read out by fluorescence of the TF fusion proteins (see below, Fig. 2a). Fusion of fluorescent proteins to PU.1 and GATA1 therefore did not alter their expression. In addition, we could not detect changes in the stability of PU.1–eYFP and GATA1–mCherry by quantitative immunostaining (Fig. 1o). Thus, although only surrogate reporters, fluorescent protein fusions can be used²¹ to quantify expression of PU.1 and GATA1 proteins in living HSPCs.

As expected, MEPs express high levels of GATA1–mCherry, but only low levels of PU.1–eYFP. GMPs express high levels of PU.1 eYFP, whereas most are negative for GATA1–mCherry (Fig. 2a). These expression patterns are identical to those previously described for endogenous TFs^{13,22}, and are similar, but not the same, as for other reporters²³. GATA1 is known to have a role downstream of GMPs²⁴, and a small GMP subpopulation co-expressed PU.1 and GATA1 (Fig. 2a).

i, j, MeGE or GM lineage choice is not changed in competitive transplantation assays (Wilcoxon rank sum test, $P=0.97$ and 0.84 , respectively. Compare Extended Data Fig. 2). **k–n**, Immunostaining for PU.1 and eYFP (**k, m**) and GATA1 and mCherry (**l, n**), day 7 of HSC differentiation. Representative examples from three independent experiments. DAPI, nuclei stain; a.u., arbitrary units. Scale bar, $10\mu\text{m}$. **o**, Normal stability of TF fusion proteins. PU.1–eYFP (4 independent experiments) or GATA1–mCherry (5 independent experiments) protein decay after $50\mu\text{M}$ Cycloheximide treatment of GMPs or preMeGE cells, respectively. Data from quantitative immunostaining against PU.1 or GATA1.

These cells do not have CMP (GEMMeg) lineage potential, and their strong GM bias suggests that GATA1 does not have a role in the GM versus MegE lineage decision of these cells (data not shown). PU.1^{mid}GATA1^{mid} progenitors mostly had only MegE potential (Fig. 2b). PU.1⁺GATA1^{hi} cells are more mature and no longer have colony potential (Fig. 2b).

HSPCs express both *PU.1* and *Gata1* mRNA before MegE versus GM lineage choice^{8,10,11}. We therefore expected CD34⁺CD16/32[−]c-Kit⁺Sca-1[−]lineage[−] CMPs¹³ to co-express both TFs. However, the vast majority express only high PU.1 levels, or GATA1 with low or no PU.1 expression. These are already committed to the GM and MegE lineage, respectively (Fig. 2c). Thus, this ‘CMP’ population is in fact a mixture of already committed GMPs and MEPs²². Haematopoietic stem cells (HSCs) already express intermediate levels of PU.1–eYFP, but no GATA1–mCherry (Fig. 2a). To identify the expected HSPC population

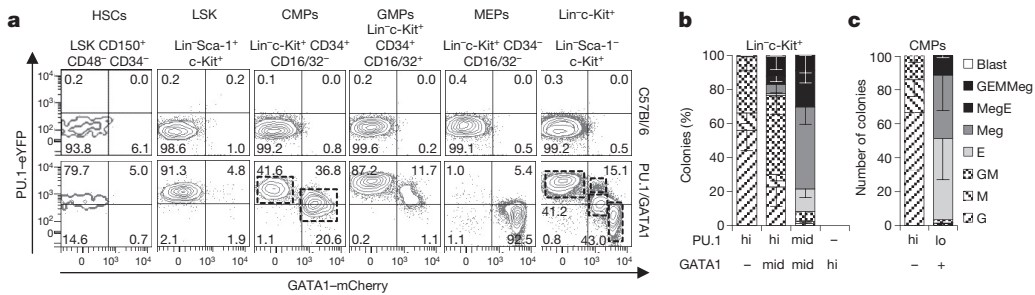


Figure 2 | PU.1-eYFP and GATA1-mCherry expression in different HSPCs. a, Flow cytometry analysis of adult bone marrow HSPCs (representative example of 4 independent experiments). **b, c**, Colony formation (mean \pm s.d.) of PU.1-eYFP^{hi}GATA1-mCherry⁺ (clonogenicity 39.6% \pm 7.1%; 4 independent experiments), PU.1-eYFP^{hi}GATA1-

mCherry^{mid} (32.2% ± 11.4%; 4), PU.1-eYFP^{mid}GATA1-mCherry^{mid} (37.1% ± 11.9%; 4) and PU.1-eYFP⁺GATA1-mCherry^{hi} (0%; 3) Lin⁻c-Kit⁺ cells (**b**), and PU.1-eYFP⁺GATA1-mCherry⁻ (58.4% ± 14.4%; 3) and GATA1-mCherry⁺ CMPs (46.3% ± 10.4%; 3) (**c**). Populations from dashed boxes in **a**.

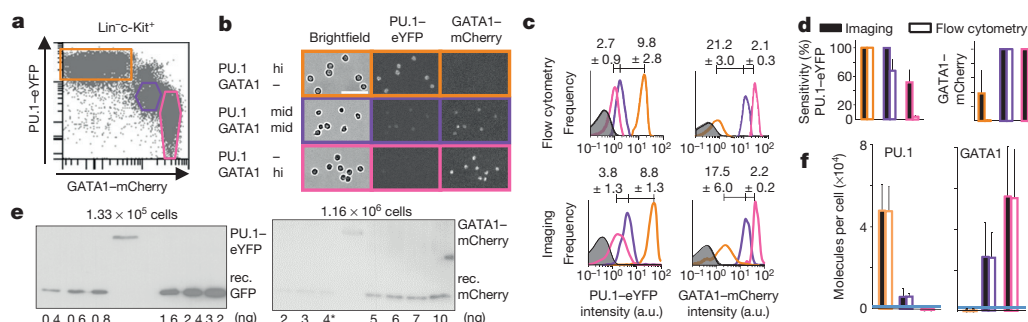


Figure 3 | Sensitive quantification of PU.1-eYFP and GATA1-mCherry protein numbers by live-cell imaging. **a, b**, HSPCs were sorted and imaged. Scale bar, 50 μ m. **c**, Relative fluorescence quantification by imaging or flow cytometry yields comparable results (mean \pm s.d., representative examples of 3 independent experiments). Filled histograms: non-fluorescent control cells. **d**, Better sensitivity of imaging over flow cytometry (mean percentages \pm s.d. of cells gated in **a** above the detection threshold). **e**, Quantification of molecule numbers in sorted PU.1-eYFP

bone marrow Lin⁻c-Kit⁺ cells and E14.5 GATA1-mCherry⁺ fetal liver cells by comparison to defined amounts of recombinant (rec.) eGFP and mCherry, respectively. Representative examples from 3 independent experiments are shown. Asterisk (*) indicates incomplete loading of sample. **f**, Estimation of molecule numbers in the populations from **a**. Mean protein abundance per cell is shown. Error bars include uncertainty from western blot quantification and fold changes from flow cytometry or imaging (mean \pm s.d.). Blue lines, imaging detection thresholds.

with GEMMeg potential downstream of PU.1^{mid}GATA1⁻ HSCs and upstream of already committed CMPs, GMPs or MEPs, we analysed the whole lineage⁻Sca-1⁺c-Kit⁺ (LSK) progenitor population (Fig. 2a). As in a previously described GATA1 reporter mouse²³, only a small number of LSK cells expressed GATA1-mCherry. The earliest HSPC population in which a small but robust GATA1-mCherry⁺ subpopulation could be detected was the MPP2 (LSK CD34⁺CD48⁺CD150⁺CD135⁻) population^{25,26} (Extended Data Fig. 5f). However, this GATA1⁺ MPP2 subpopulation is already MegE committed (see below).

The PU.1-GATA1 switch model is based on near-stoichiometric PU.1 and GATA1 co-expression, and thus mutual functional inhibition, before GM versus MegE lineage decision^{6,7}. In contrast, we could not identify a PU.1-eYFP⁺GATA1-mCherry⁺ HSPC population with robust GEMMeg potential. However, this data from snapshot cytometry analysis cannot exclude that differentiating PU.1^{mid}GATA1⁻ HSCs may quickly pass through a state with similar PU.1 and GATA1 expression. We therefore extended approaches for long-term imaging and single-cell tracking^{21,27,28} to allow continuous live quantification of PU.1-eYFP and GATA1-mCherry in differentiating HSPCs²⁹. PU.1-eYFP and GATA1-mCherry fluorescence was better detected by imaging than by flow cytometry, with a greater dynamic range and higher sensitivity (Fig. 3a-d). TF protein numbers in individual cells were estimated by comparison to defined amounts of recombinant fluorescent proteins in western blot analyses (Fig. 3e, f and Extended Data Table 2; for gel source data, see Supplementary Fig. 1).

We cultured HSCs under conditions allowing both MegE and GM differentiation, which were detected by expression of GATA1-mCherry or CD16/32 (ref. 13), respectively (Extended Data Fig. 6). Due to its overlapping expression, PU.1-eYFP expression alone does not allow early lineage detection (Extended Data Fig. 6a). Although these culture conditions do not resemble all possible *in vivo* conditions,

they do allow differentiation into all relevant lineages, thus enabling analysis of the core mechanisms expected to underlie MegE versus GM lineage choice.

We quantified absolute PU.1 and GATA1 protein levels in single differentiating HSCs and their progeny, throughout up to 11 generations (Fig. 4a, Extended Data Fig. 7). About 6.5×10^6 total measurements, including 3.7×10^5 (1.8×10^5 manually curated) fluorescence measurements, in four experiments with 1,080 CD16/32 and 681 GATA1-mCherry onsets from 256 different HSC colonies were analysed (Extended Data Table 3). As expected, cells differentiating into the GM lineage had increased PU.1-eYFP levels over time and later expressed CD16/32 (Fig. 4b, left panels, Supplementary Video 1). Unexpectedly, we did not detect GATA1-mCherry expression at any point during GM differentiation. This is in contrast to expectations from the previous model in which the PU.1-GATA1 switch acts as the initiator of this lineage choice. Our detection limit for GATA1-mCherry is about 1,900 molecules per cell. In about half of all GM differentiations, PU.1-eYFP levels steadily increased from the starting HSC until the onset of CD16/32 expression. In the other half, PU.1-eYFP levels transiently dropped, then steadily increased. However, only $25 \pm 5\%$ of all GM time-course analyses showed PU.1 numbers dropping below 8,100 molecules (the average expression in HSCs). Moreover, only about $1 \pm 1\%$ of GM-differentiating cells transiently dropped to below 2,000 PU.1 molecules, and thus to similar levels of potentially maximally expressed GATA1 molecules. GATA1 levels thus do not have a relevant role during GM differentiation.

Cells that expressed detectable GATA1-mCherry, during HSC differentiation or in freshly sorted GATA1⁺ MPP2 cells, always further differentiated into PU.1-eYFP⁻GATA1-mCherry⁺ MegE cells (Fig. 4b, right panels, Supplementary Video 2). This confirms that GATA1 expression onset is a marker of MegE lineage commitment (Extended

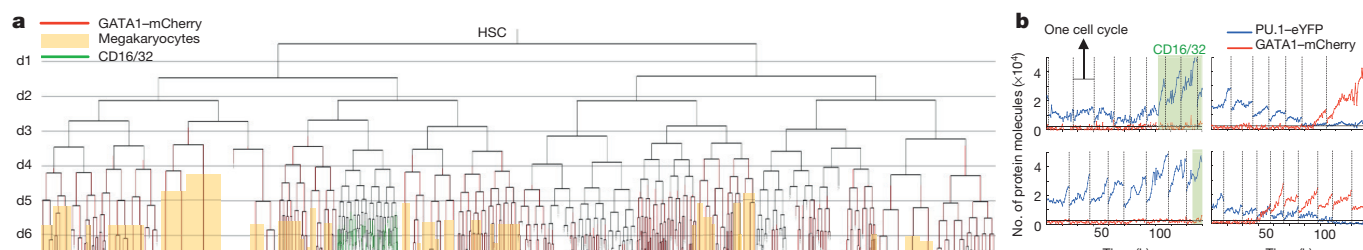


Figure 4 | Single-cell dynamics refute random PU.1:GATA1 ratios as initiators of early myeloid lineage choice. **a**, Single-cell genealogy of a differentiating HSC. CD16/32 detection by live antibody staining, megakaryocytes determined by cell morphology. **b**, Typical expression

dynamics of PU.1 and GATA1 expression of GM (left panels) or MegE differentiating cells (right panels). Black horizontal lines represent detection threshold for GATA1-mCherry.

Data Fig. 6). Importantly, this was independent of PU.1 levels during GATA1 expression onset, demonstrating that PU.1:GATA1 ratios do not regulate MegE commitment. In addition, PU.1-eYFP levels already decreased before detectable GATA1-mCherry expression in $63 \pm 19\%$ of cases (compare Fig. 4b, upper right panel). Thus, GATA1 expression is not the cause of PU.1 downregulation. In all remaining cases, GATA1 was upregulated, while PU.1 was still expressed at different levels. However, in these cases, cells later always differentiated into a PU.1⁺GATA1^{high} MegE state, showing that different PU.1 levels during lineage decision-making are irrelevant for the onset of GATA1 expression.

In conclusion, we did not observe a reproducible PU.1–GATA1 double-positive stage through which all differentiating HSCs pass (Supplementary Video 3). PU.1 and GATA1 are independently regulated at the start of GM or MegE differentiation. These observed protein dynamics are incompatible with random and cross-regulatory PU.1–GATA1 co-expression acting as the central mechanism that initiates MegE versus GM lineage choice³. This conclusion is in line with observations that lineage choice is still possible after deletion of PU.1 or GATA1, where only further maturation of committed cells is impaired^{1,2,30}. Our data are compatible with that from other reporter mice²³. However, our results also demonstrate discrepancies between protein and mRNA expression in uncommitted HSPCs^{8,31}, the latter of which had originally led to the development of currently accepted models.

Although we demonstrate that ratios of total PU.1:GATA1 protein numbers are not the central mechanism that initiates HSPC lineage decisions, we cannot exclude the possibility that only a very small subset of expressed PU.1 proteins may be actively competing with potentially existing GATA1 protein expressed below our detection limit. The PU.1–GATA1 switch may be involved in the lineage choice of other cell types²⁴ not analysed here. Likewise, other TF switches could be involved in GM versus MegE or other lineage choices. We conclude that physical PU.1–GATA1 interaction and antagonism^{6,7} could serve as an execution and/or reinforcing mechanism making terminal differentiation irreversible, but not as a decision-making mechanism inducing it. We expect other TFs and signalling pathways activated by extracellular signals to be upstream regulators of lineage-specific TFs, and their complex interplay will be of interest for future analyses.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 June 2015; accepted 13 May 2016.

1. Pevny, L. *et al.* Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* **349**, 257–260 (1991).
2. Scott, E. W., Simon, M. C., Anastasi, J. & Singh, H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* **265**, 1573–1577 (1994).
3. Graf, T. & Enver, T. Forcing cells to change lineages. *Nature* **462**, 587–594 (2009).
4. Tsai, S. F., Strauss, E. & Orkin, S. H. Functional analysis and *in vivo* footprinting implicate the erythroid transcription factor GATA-1 as a positive regulator of its own promoter. *Genes Dev.* **5**, 919–931 (1991).
5. Chen, H. *et al.* PU.1 (Spi-1) autoregulates its expression in myeloid cells. *Oncogene* **11**, 1549–1560 (1995).
6. Zhang, P. *et al.* PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding. *Blood* **96**, 2641–2648 (2000).
7. Nerlov, C., Querfurth, E., Kulesha, H. & Graf, T. GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. *Blood* **95**, 2543–2551 (2000).
8. Miyamoto, T. *et al.* Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. *Dev. Cell* **3**, 137–147 (2002).
9. Huang, S., Guo, Y.-P., May, G. & Enver, T. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Dev. Biol.* **305**, 695–713 (2007).
10. Akashi, K. *et al.* Transcriptional accessibility for genes of multiple tissues and hematopoietic lineages is hierarchically controlled during early hematopoiesis. *Blood* **101**, 383–389 (2003).

11. Månsson, R. *et al.* Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity* **26**, 407–419 (2007).
12. Hoppe, P. S., Coutu, D. L. & Schroeder, T. Single-cell technologies sharpen up mammalian stem cell research. *Nat. Cell Biol.* **16**, 919–927 (2014).
13. Akashi, K., Traver, D., Miyamoto, T. & Weissman, I. L. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404**, 193–197 (2000).
14. Orkin, S. H. & Zon, L. I. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**, 631–644 (2008).
15. Etzrodt, M., Ende, M. & Schroeder, T. Quantitative single-cell approaches to stem cell research. *Cell Stem Cell* **15**, 546–558 (2014).
16. Kueh, H. Y., Champhekar, A., Nutt, S. L., Elowitz, M. B. & Rothenberg, E. V. Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. *Science* **341**, 670–673 (2013).
17. Kirstetter, P., Anderson, K., Porse, B. T., Jacobsen, S. E. W. & Nerlov, C. Activation of the canonical Wnt pathway leads to loss of hematopoietic stem cell repopulation and multilineage differentiation block. *Nat. Immunol.* **7**, 1048–1056 (2006).
18. Rosenbauer, F. *et al.* Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat. Genet.* **36**, 624–630 (2004).
19. Moreau-Gachelin, F. *et al.* Spi-1/PU.1 transgenic mice develop multistep erythroleukemias. *Mol. Cell Biol.* **16**, 2453–2463 (1996).
20. Heyworth, C., Pearson, S., May, G. & Enver, T. Transcription factor-mediated lineage switching reveals plasticity in primary committed progenitor cells. *EMBO J.* **21**, 3770–3781 (2002).
21. Filipczyk, A. *et al.* Network plasticity of pluripotency transcription factors in embryonic stem cells. *Nat. Cell Biol.* **17**, 1235–1246 (2015).
22. Pronk, C. J. H. *et al.* Elucidation of the phenotypic, functional, and molecular topography of a myeloerythroid progenitor cell hierarchy. *Cell Stem Cell* **1**, 428–442 (2007).
23. Arinobu, Y. *et al.* Reciprocal activation of GATA-1 and PU.1 marks initial specification of hematopoietic stem cells into myeloerythroid and myelolymphoid lineages. *Cell Stem Cell* **1**, 416–427 (2007).
24. Iwasaki, H. *et al.* Identification of eosinophil lineage-committed progenitors in the murine bone marrow. *J. Exp. Med.* **201**, 1891–1897 (2005).
25. Miyawaki, K. *et al.* CD41 marks the initial myelo-erythroid lineage specification in adult mouse hematopoiesis: redefinition of murine common myeloid progenitor. *Stem Cells* **33**, 976–987 (2015).
26. Pietras, E. M. *et al.* Functionally distinct subsets of lineage-biased multipotent progenitors control blood production in normal and regenerative conditions. *Cell Stem Cell* **17**, 35–46 (2015).
27. Rieger, M. A., Hoppe, P. S., Smejkal, B. M., Eitelhuber, A. C. & Schroeder, T. Hematopoietic cytokines can instruct lineage choice. *Science* **325**, 217–218 (2009).
28. Eilken, H. M., Nishikawa, S. & Schroeder, T. Continuous single-cell imaging of blood generation from haemogenic endothelium. *Nature* **457**, 896–900 (2009).
29. Hilsenbeck, O. *et al.* Software tools for single-cell tracking and quantification of cellular and molecular properties. *Nat. Biotechnol.* <http://dx.doi.org/10.1038/nbt.3626> (2016).
30. Mancini, E. *et al.* FOG-1 and GATA-1 act sequentially to specify definitive megakaryocytic and erythroid progenitors. *EMBO J.* **31**, 351–365 (2012).
31. Hu, M. *et al.* Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev.* **11**, 774–785 (1997).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to S. Ammersdoerfer, H. Oller, C. Raithel, B. Vogel and A. Ziegler for technical support. This work was supported by an EMBO long-term fellowship to M.E., the ‘EUCOMM: Tools for Functional Annotation of the Mouse Genome’ (EUCOMMTOOLS) project (FP7-HEALTH-F4-2010-261492) to A.B., the ERC (starting grant Latent Causes), the German Federal Ministry of Education and Research (BMBF), the German Research Foundation (DFG) within the SPPs 1395 (InKoMBio) and 1356 to F.J.T., and DFG SFB 684 to T.S. and the SNF to T.S. T.S. and O.H. acknowledge financial support from SystemsX.ch.

Author Contributions P.S.H. planned and performed experiments and analysed data; M.Sc. programmed and applied quantitative imaging software and performed protein quantification and statistical analysis with M.St., C.M. and F.J.T. M.Sc., D.L., K.D.K., M.En., N.M., M.A.R., N.A., M.E., and A.F. provided support for time-lapse imaging, flow cytometry and software development. O.H. and B.S. programmed single-cell tracking software with T.S. D.L.C. contributed to immunofluorescence staining. I.B., H.L. and A.B. contributed to generation of GATA1^{mCherry} mice. O.E., A.G. and C.N. provided the PU.1^{eYFP} mouse and competitive transplantations. F.J.T. designed and supervised the data analysis and modelling part. T.S. designed the study, programmed software, analysed data and wrote the paper with P.S.H. All authors read and commented on the final manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.S. (tim.schroeder@bss.ethz.ch).

METHODS

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Generation of *Gata1-mCherry* knock-in mice. The knock-in construct was cloned by using conventional restriction-enzyme-mediated cloning and recombineering³² using the BAC RPCIB731C02198Q (Source BioScience) that contained the *Gata1* locus. The final knock-in construct consisted of a 5.0 kbp 5'-homology arm lasting until the last codon of *Gata1* (skipping the endogenous stop-codon) followed by a short linker sequence (5'-AGAGCATCAGGTACAGTGGAGCT-3'), the coding sequence for mCherry³³ and a FRT (FLP recognition target)-flanked phospho-glycerate kinase (PGK) promoter-driven neomycin (neo) resistance gene and the 4.6 kbp 3'-homology arm. After removal of the neo selection marker, the *Gata1-mCherry* fusion mRNA transcript utilizes the endogenous 3' UTR.

JM8 mouse embryonic stem (ES) cell lines derived from the C57Bl/6N strain were grown on gelatinized tissue culture plates. Cells were maintained in knockout DMEM media (Gibco) supplemented with 2 mM glutamine, 1% β -mercaptoethanol (360 μ l in 500 ml PBS, sterile-filtrated), 10–15% fetal calf serum (Invitrogen) and 500 U ml⁻¹ ESGRO leukaemia-inhibitory factor (Millipore). Electroporation of ES cells was carried out in a 25-well cuvette using the ECM 630 96-well electroporator / HT-200 automatic plate handler (BTX Harvard Apparatus; set at 700 V, 400 Ω , 25 μ F). Immediately before electroporation, cell suspensions of $\sim 10^7$ cells and ~ 2.5 μ g of linearized targeting vector DNA were mixed in a final volume of 120 μ l PBS. Cells were seeded onto a gelatinized 10-cm dish and colonies were picked after 8–9 days of puromycin (3 μ g ml⁻¹) selection. The colonies were expanded in four copies of 96-well plates for archiving and characterization. Cells were frozen in supplemented knockout DMEM with 10% DMSO and stored in vapour over liquid nitrogen. After identification of positive clones, cells were thawed and expanded for aggregation.

Correctly targeted ES cell clones were identified by Southern blot using probes at the designated locations (Extended Data Fig. 1a) after digestion of genomic DNA with the restriction enzymes BamHI and XbaI, respectively. PCR primers for generating the Southern probes from BAC DNA were 5'-CAGCCACTG CCAAAATAGGTGGAG-3' and 5'-CTCCACCTATTGGGCAGTGGCTG-3' (5'-probe) and 5'-CTGAAGTGGTCTCTGGACTTAC-3' and 5'-TGAGGAAGA GGAAGGATGTGAAG-3' (3'-probe).

From one ES cell clone, germline chimaeras were generated by ES-cell aggregation with CD1 morulae and the FRT-flanked neo selection cassette was deleted *in vivo* by a FLP-e deleter strain by recombinase-mediated excision³⁴.

Animal experiments were approved by veterinary office of Canton Basel-Stadt, Switzerland and Regierung von Oberbayern.

Genotyping. PCR primers for checking presence or absence of the NEO cassette were 5'-GCATGGACGAGCTGTACAAG-3', 5'-CTGCACGAGAC TAGTGAGAC-3' and 5'-GCAGGAGAATGGGAAATGTG-3' leading to a 223 bp band after successful removal. Unsuccessful removal would have led to a 387 bp band. Primers for checking the presence or absence of FLP-recombinase were 5'-GTTCTATATGCTGCCACTCC-3' and 5'-GAGCGATAAGCGTGCTTCTG-3' leading to 176 bp band at its presence. *GATA1^{mCherry}* mice were genotyped using the primers 5'-GCATGGACGAGCTGTACAAG-3', 5'-AGGTACTGCCC ACCTCTATC-3' and 5'-GCAGGAGAATGGGAAATGTG-3' leading to a 297 bp band in the case of wild-type *Gata1* and a 223 bp band in the case of *Gata1-mCherry*.

Isolation and staining of primary HSPCs and blood. Male and female mice for blood counts, bone marrow analysis, and time-lapse movies were killed at the age of 12–16 weeks. Blood counts were quantified on an Abc Animal Blood Counter (scil animal care company). Isolation of primary cells and flow cytometry sorting was performed as described^{13,22,35,36}. All flow cytometry was performed on a FACSAria I or III (BD Bioscience). In brief, pelvis, femurs, tibiae, humeri and vertebrae of adult mice were isolated, crushed and incubated with anti-CD16/32 antibody (clone 2.4G2, BD Pharmingen, or clone 93, eBioscience) before staining with the desired antibody cocktail. Cells prepared for sorting HSCs were subjected to ACK Lysing Buffer (Lonza) after crushing, followed by lineage depletion using biotinylated antibodies against CD3e (clone 145-2C11), CD11b (clone M1/70), CD19 (clone 1D3), CD41 (clone MWReg30), B220 (clone RA3-6B2), Gr-1 (clone RB6-8C5) and TER-119 (clone TER-119, all eBioscience) and streptavidin-conjugated beads Roti-MagBeads (Carl Roth). The following antibodies were used for staining: anti-CD34 (RAM34), anti-CD48 (HM48-1), anti-CD105 (MJ7/18), anti-CD117 (2B8), anti-CD117 (ACK2), anti-CD135 (A2F10), anti-Sca-1 (D7, all eBioscience) and anti-CD150 (TC15-12F12.2, BioLegend). Different HSPC types within the lineage⁻Sca-1⁺c-Kit⁺ (LSK) population (Fig. 1b) were identified as: HSC, LSK CD150⁺CD34⁺CD48⁻; MPP1, LSK CD150⁺CD34⁺CD48⁻; MPP2, LSK CD150⁺CD34⁺CD48⁺; MPP3/4, LSK CD150⁺CD34⁺CD48⁺. When cells were prepared for sorting myeloid progenitors, CD41-biotin was omitted. CD41 was also omitted for the analysis of the MPP subpopulations (Extended Data Fig. 5f).

Single-cells were sorted into 384-well plates (Greiner Bio-One), according to manufacturer's instructions, on a FACSAriaIII (BD Bioscience).

E14.5 fetal livers were isolated, individualized, and stained with antibodies for analysis. For GATA1 protein numbers quantification, GATA1-mCherry⁺ cells were sorted.

Colony assays. All colony assays were performed in Methocult GF M3434 (STEMCELL Technologies) according to manufacturer's instructions.

Competitive transplantations. After ACK Lysis (Life Technologies), freshly isolated bone marrow from CD45.1 homozygous C57Bl/6 wild-type mice was mixed 1:1 with bone marrow from either CD45.2 homozygous C57Bl/6 wild-type mice or CD45.2 homozygous PU.1^{eYFP}GATA1^{mCherry} (P/G) mice. Cells were frozen in 90% IMDM (Life Technologies) and 10% DMSO (Sigma Aldrich) and stored above liquid nitrogen until further usage. For transplantations, cells were thawed, counted and 10⁶ living cells were transplanted into the tail vein of lethally irradiated CD45.1/CD45.2 heterozygous wild-type mice. Lineage contribution of donor cell mixtures was determined by bone marrow collection and staining²² after 6–7 weeks and plotted as the ratio of preMegE/preGM cells (Fig. 1i) and whole MegE/whole GM (Fig. 1l) lineage cells of wild-type:wild-type and P/G:wild-type donor mixes.

Immunostaining. Immunostaining was performed after permeabilization with 0.2% Triton-X (Applichem) with 8 μ g ml⁻¹ anti-PU.1 (T-21), 8 μ g ml⁻¹ anti-GATA1 (N6) (both Santa Cruz), 10 μ g ml⁻¹ anti-GFP (Aves Labs) and 5 μ g ml⁻¹ anti-mCherry (ab167453) (Abcam) primary antibodies in 10% donkey serum in TBS-T (Tris-buffered saline, 0.1% Tween 20) overnight at 4 °C, three washing steps of each 5 min and 10 μ g ml⁻¹ Alexa Fluor dyes conjugated donkey secondary antibodies (Jackson ImmunoResearch) for 1 h at room temperature in 10% donkey serum in TBS-T. Images were acquired on a Nikon Eclipse Ti-E microscope. Fluorescent signals were quantified by segmentation of nuclear DAPI staining and background subtraction (without 'gain') as described³⁷. In order to compare transcription factors and their fusions regarding their stability, cells were kept in medium additionally supplied with 50 μ M Cycloheximide and split into several vials. For immunostainings, cells were transferred to poly-L-lysine (Sigma Aldrich) coated object slides, stored for 10 min at 4 °C, fixed with paraformaldehyde (Sigma Aldrich), stored at 4 °C and stained (see above).

Western blot analysis. For western blotting, cells were spun down at designated time points and directly lysed in Laemmli-buffer, boiled at 100 °C and frozen at -20 °C until further usage in SDS-PAGE analysis. For protein number quantification, designated amounts of recombinant GFP (Clontech) and mCherry (ChromoTek) were used. Gels were run on a 10% SDS-gel, blotted onto a PVDF-membrane (BioRad) and incubated with either anti-GFP (7.1, 13.1; Roche) or anti-mCherry antibody (1C51; abcam). Chemiluminescence was detected using horseradish peroxidase (HRP)-conjugated secondary antibodies (Jackson), ECL Plus Western Blotting Detection Reagents (GE Healthcare) and medical X-ray films (Fujifilm). Signal intensities were quantified using ImageJ software.

Cytospins. Cells were spun on object slides as per the manufacturer's instructions (Hettich), dried and stained with May-Grünwald (Carl Roth) and Giemsa solution (Sigma Aldrich).

Time-lapse image acquisition and tracking. Time-lapse imaging was performed at 37 °C in fibronectin (Takara Bio) coated channel slides μ -slide VI^{0.4} (ibidi), in StemSpan SFEM (STEMCELL Technologies) medium supplemented with 100 ng ml⁻¹ SCF, 100 ng ml⁻¹ TPO, 10 ng ml⁻¹ IL-3, 10 ng ml⁻¹ IL-6 (all mouse; PeproTech), 5 U ml⁻¹ EPO (human; PromoKine), 50 U ml⁻¹ penicillin, 50 μ g ml⁻¹ Streptomycin (Invitrogen), self-labelled Alexa Fluor 647 (Invitrogen) anti-CD16/32 antibody (2.4G2) and 5% CO₂ using an Axio Observer Z1 microscope (Zeiss). A HXP 120 (Zeiss) was used as fluorescent light source. 46HE, 43HE (both Zeiss) and Cy5 (AHF) filter sets were used to detect eYFP, mCherry and Alexa Fluor 647, respectively, at exposure times between 400–1,500 ms using an AxioCam HRm (Zeiss). Brightfield pictures were acquired every 60–120 s, fluorescent pictures for the quantification of PU.1-eYFP and GATA1-mCherry were acquired every 30 min and every 3–4 h for the detection of CD16/32-Alexa-Fluor-647. Pictures used for quantifications were saved in lossless TIF or PNG format. Single-cell tracking and image quantification were performed using self-written software as described^{21,27–29}.

Inference of absolute protein numbers. *Western blot dilution assays.* For protein number quantification, known cell numbers of PU.1eYFP^{hi} progenitors and GATA1-mCherry⁺ E14.5 fetal livers cells were resolved by western blotting on 10% polyacrylamide gels and compared with different levels of recombinant GFP protein (Clontech) or mCherry (Antibodies Online). PU.1-eYFP, GATA1-mCherry, GFP and mCherry proteins were detected using an anti-GFP primary antibody (Roche) or an anti-mCherry antibody (Abcam). All quantifications of band intensities were performed using ImageJ software by manually drawing a gate around the bands and subtracting the mean of the same area above and below the bands for primary HSPCs.

Model. After comparing several models to describe the data, we found a sigmoidal function to best describe the relationship between the dilution of the standard x and the observed intensity y :

$$y(x) = \left(\frac{\lambda x^n}{K^n + x^n} \right) \cdot \epsilon \quad (1)$$

Here, the exponent n determines the steepness of the sigmoidal, K sets the inflection point, λ is the maximum of the curve and ϵ is a lognormally distributed error term with expectation 1 and standard deviation σ as suggested for western blot data³⁸. This model outperformed other models, such as linear models with and without intercept according to the Bayesian information criterion and coefficient of variation between replicates. We estimate model parameters using 10,000 local optimizations initialized according to Latin-hypercube sampling.

We determine the number of proteins, P_j , per cell from the sample intensity, y_j , of replicate j (that is, western blot) as

$$P_j = \frac{K_j}{\left(\frac{\lambda_j}{y_j} - 1 \right)^{\frac{1}{n_j}}} \cdot \frac{1}{c_j \cdot w} \quad (2)$$

where c_j is number of loaded cells and w is the molecular weight for the protein of interest. The first term is obtained by solving equation (1) for x , whereas the second term relates dilution in nanogram to absolute protein numbers per cell. The parameters λ_j , K_j and n_j have been obtained by the local optimization.

Error propagation. As P_j is a combination of uncertain variables, we obtained error bars for each P_j individually by applying standard error propagation to account for uncertainties in the number of cells c_j (we assume a standard deviation of 10%) and uncertainties in the model (estimated via the standard deviation σ of our noise model ϵ_j in equation (1)). However, we find that the uncertainties for each individual replicate P_j are always with a factor <0.3 smaller than the inter-replicate standard deviations. Therefore, we only consider the standard deviation across replicates, as this is the dominant source of uncertainty in our procedure.

Mapping protein numbers to different HSPC populations. Inference. From reference cell types used for western blot dilution assays, we map mean protein numbers to all other cell types using the mean fluorescence intensities from flow cytometry. As we observe fluorescence intensities from fusion protein levels, we assume a linear relation between flow cytometry intensities and molecule numbers. We obtain the average protein amount per flow cytometry intensity of our reference cell type r as $\psi_r = \frac{\bar{P}^r}{\overline{\text{MFI}}^r}$. Here, $\overline{\text{MFI}}^r$ is the mean fluorescence intensity of our protein of interest in the reference population (for example, PU.1-eYFP^{hi}) used for the western blotting, averaged over N replicates: $\overline{\text{MFI}}^r = \frac{1}{N} \sum_{i=1}^N \text{MFI}_i^r$. By \bar{P}^r we denote the estimated amount of protein from western blotting via equation (2), averaged over three replicates j .

We calculate the average amount of protein in any other population of interest (for example, GMPs) as

$$\bar{P}^{\text{GMP}} = \psi_r \cdot \overline{\text{MFI}}^{\text{GMP}} = \frac{\overline{\text{MFI}}^{\text{GMP}}}{\overline{\text{MFI}}^r} \cdot \bar{P}^r,$$

where $\overline{\text{MFI}}^{\text{GMP}}$ is the mean fluorescence intensity of our protein of interest in the GMP population averaged over N replicates.

Error Propagation. To obtain error bars for the protein amount in a population of interest, we perform error propagation, taking into account the uncertainty ΔP^r in protein numbers (as described above, we only consider inter-replicate variation), as well as uncertainty in the mean fluorescence intensities (standard deviation ΔMFI^r over the MFI_i^r) (for example, GMPs):

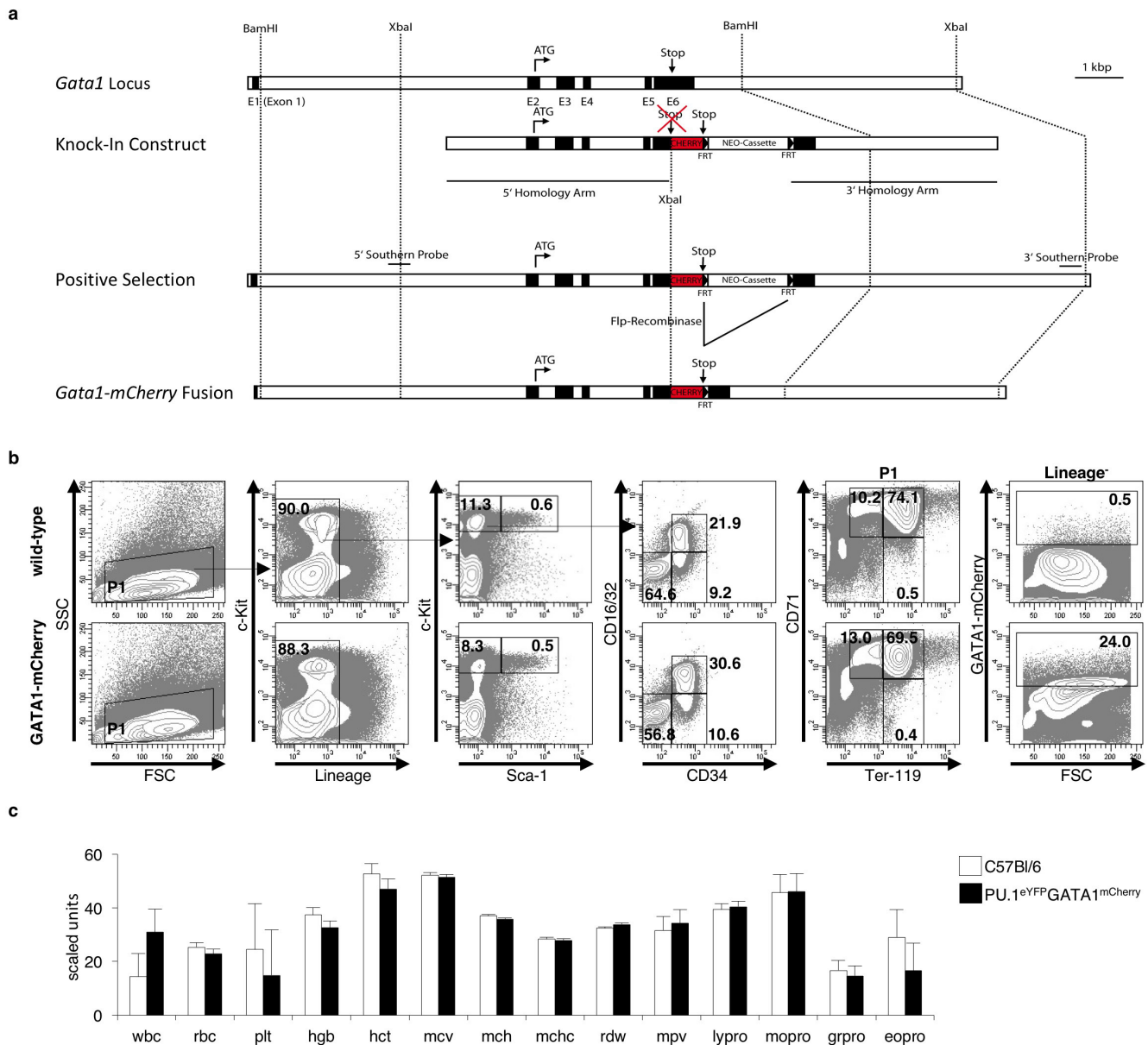
$$\Delta P^{\text{GMP}} = \sqrt{(\Delta \text{MFI}^{\text{GMP}})^2 \cdot \left(\frac{\bar{P}^r}{\overline{\text{MFI}}^r} \right)^2 + (\Delta \text{MFI}^r)^2 \cdot \left(\frac{\overline{\text{MFI}}^{\text{GMP}} \cdot \bar{P}^r}{(\overline{\text{MFI}}^r)^2} \right)^2 + (\Delta P^r)^2 \cdot \left(\frac{\overline{\text{MFI}}^{\text{GMP}}}{\overline{\text{MFI}}^r} \right)^2}$$

Mapping protein numbers to pixel intensities in imaging. The mean fluorescence intensity of the first time points was used to calibrate PU.1-eYFP protein abundance in time-lapse experiments. Whenever a movie cell exceeds twice the detection limit in the GATA1-mCherry channel for more than five consecutive time points, the cell itself and all its descendants were annotated as GATA1-mCherry positive. Mean protein abundance of GATA1 positive movie cells has been calibrated to the mean protein abundance of PU.1^{mid}GATA1^{mid} in flow cytometry. Protein levels are then interpolated linearly³⁸.

Single-cell tracking and fluorescence quantification. Single-cell tracking was performed as described^{21,27–29}. Briefly, self-written software allows following individual cell identities over many days in order to generate genealogy trees. Fluorescence image normalization was applied as described³⁷. Custom written software semi-automatically identifies shapes of tracked cells and quantifies protein levels resulting in normalized intensity time traces independently of timing and location in time-lapse imaging. Detection thresholds were determined by simulating *in silico* background cells based on manually selected pixels containing only background signal and based on manually inspected cell areas. The 99% quartile of the resulting distribution of *in silico* background cells is referred to as the detection threshold, which is extrapolated to protein numbers for each movie as described above.

Code availability. Software used for single cell tracking and fluorescence quantification used in this study is published and open sourced²⁹.

32. Liu, P., Jenkins, N. A. & Copeland, N. G. A highly efficient recombineering-based method for generating conditional knockout mutations. *Genome Res.* **13**, 476–484 (2003).
33. Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* **22**, 1567–1572 (2004).
34. Dymecki, S. M. Flp recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice. *Proc. Natl Acad. Sci. USA* **93**, 6191–6196 (1996).
35. Kiel, M. J. *et al.* SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* **121**, 1109–1121 (2005).
36. Wilson, A. *et al.* Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* **135**, 1118–1129 (2008).
37. Schwarzfischer, M. *et al.* Efficient fluorescence image normalization for time lapse movies. *Proc. Microsc. Image Anal. with Appl. Biol.* (2011).
38. Kreutz, C. *et al.* An error model for protein quantification. *Bioinformatics* **23**, 2747–2753 (2007).

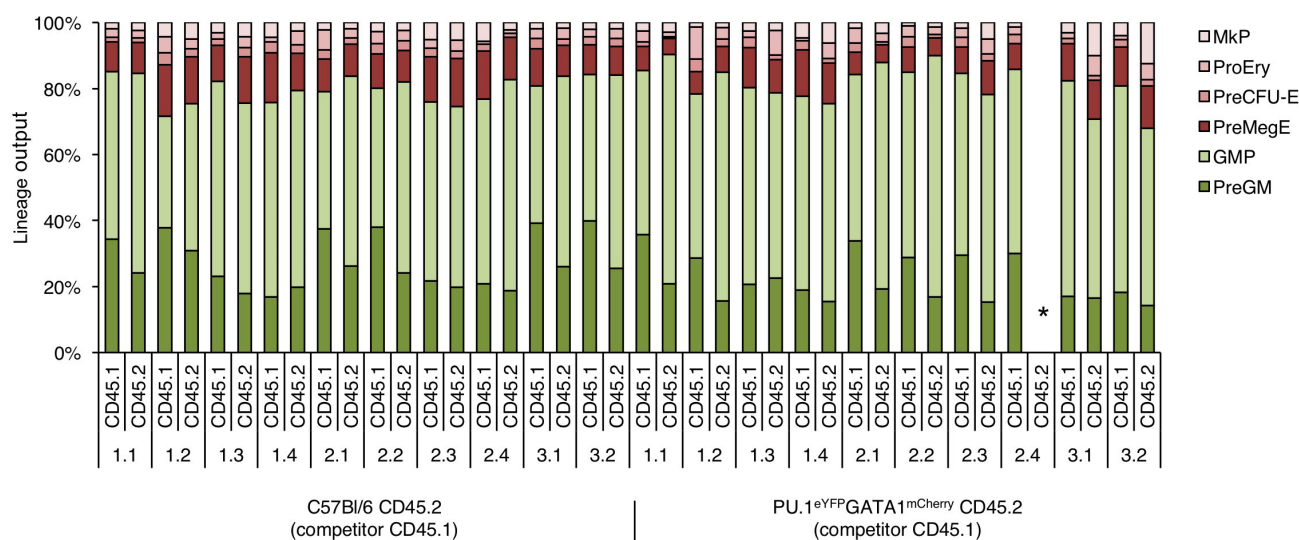


Extended Data Figure 1 | Knock-in of fluorescent proteins does not alter the composition of fetal livers and adult peripheral blood.

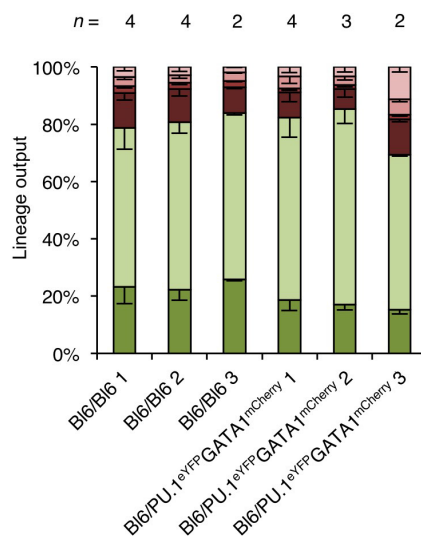
a, *Gata1-mCherry* knock-in strategy. BamHI and XbaI were chosen to generate restriction fragment length polymorphisms (from 11.1 kbp to 5.7 kbp in the case of XbaI and from 9.9 kbp to 11.1 kbp in the case of BamHI) to screen for successful knock-ins. Genomic sequences for Southern probes were identified at indicated positions. The final knock-in construct contained a 5.0 kilo base pairs (kbp) long 5' homology arm until the last codon of *Gata1*, a short linker sequence (5'-AGAGCATCAGGTACCACTGGAGCT-3'), the open reading frame (ORF) of mCherry, a FRT-flanked Neomycin-resistance cassette (including a eukaryotic and a prokaryotic promoter and a polyadenylation signal) and a 4.6 kbp long 3' homology arm. **b**, Fetal livers of E14.5 embryos were collected, subjected to Ficoll-density centrifugation,

pooled (C57Bl/6, 7 fetal livers; GATA1-mCherry, 6 fetal livers) and analysed by flow cytometry. Shown are percentages of the parental gate. **c**, Peripheral blood counts of adult mice (C57Bl/6 $n = 6$ biological replicates; PU.1^{eYFP}GATA1^{mCherry} $n = 9$; error bars, mean \pm s.d.). wbc, white blood cells (200 cells per mm³); rbc, red blood cells (4×10^5 cells per mm³); plt, platelets (20 cells per mm³); hgb, haemoglobin (0.4 g dl⁻¹); hct, haematocrit (%); mcv, mean corpuscular volume (μ m³); mch, mean corpuscular haemoglobin (0.4 pg); mchc, mean corpuscular haemoglobin concentration (g dl⁻¹); rdw, red cell distribution width (0.4%); mpv, mean platelet volume (0.2 μ m³); lypro, % lymphocytes of wbc (2%); mopro, percentage monocytes of wbc (0.1%); grpro, percentage of granulocytes of wbc (%); eopro, percentage eosinophils of wbc (0.2%). No significant difference, one-way MANOVA ($P > 0.09$).

a

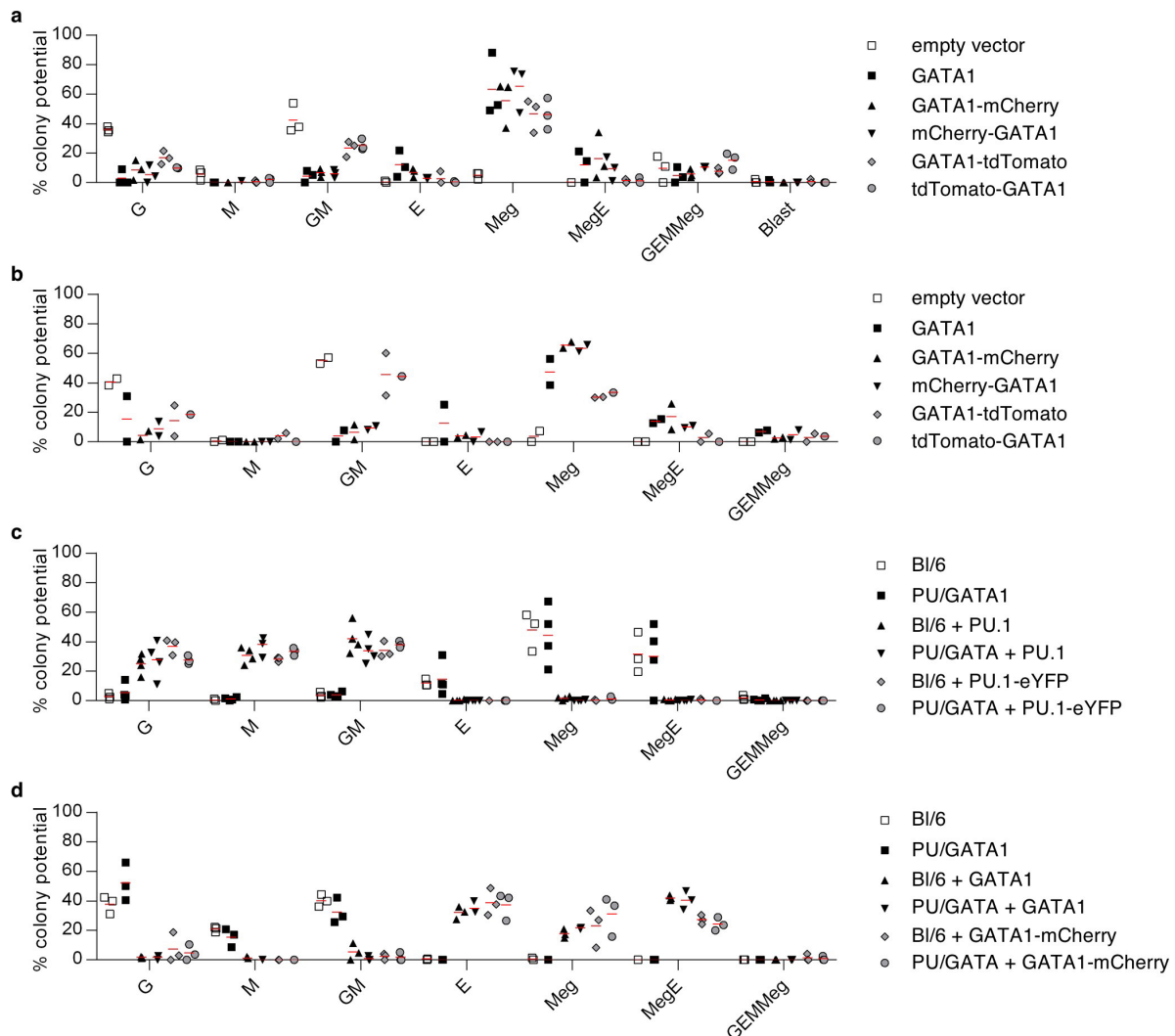


b



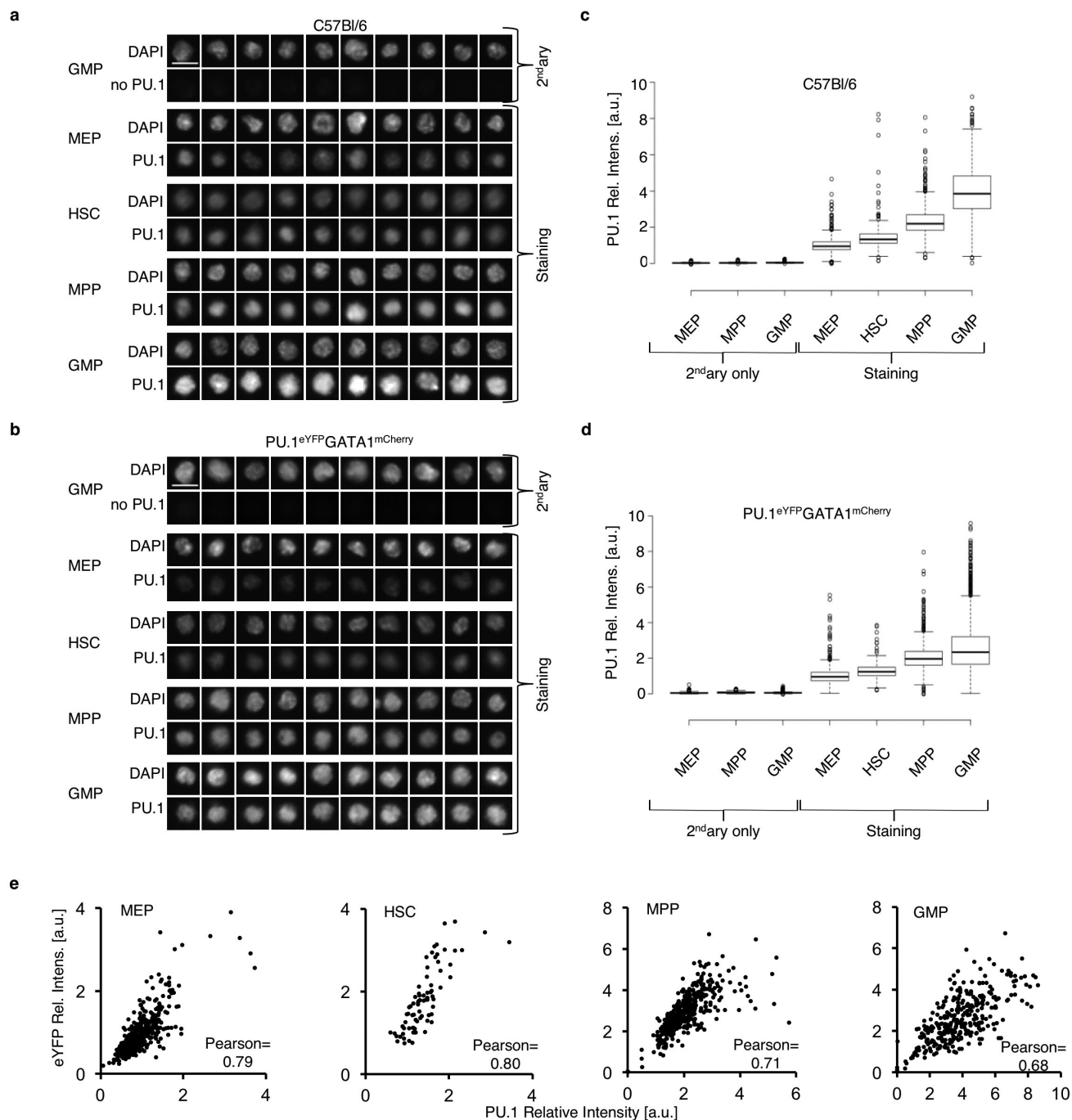
Extended Data Figure 2 | MegE versus GM-lineage differentiation of GATA1-mCherry/PU.1-eYFP cells *in vivo* is unaltered upon competitive transplantation. **a**, 10^6 bone marrow cells each from CD45.1 C57Bl/6 and CD45.2 C57Bl/6 or from CD45.1 C57Bl/6 and CD45.2 PU.1^{eYFP}GATA1^{mCherry} mice were transplanted into lethally irradiated recipient mice and bone marrow progenitor cell composition was analysed after 6–7 weeks. *X* and *Y* (*X.Y*, e.g. 1.1)

denote donor pair (*X*) and recipient mouse (*Y*). MkP, megakaryocyte progenitor; proEry, proerythroblast; preCFU-E, pre-colony-forming unit erythrocyte; preMegE, premegakaryocyte-erythrocyte progenitor; preGM, pregranulocyte-macrophage progenitor. Asterisk (*) indicates data excluded owing to low donor contribution. **b**, Summarized bone marrow lineage contribution per donor pair (from **a**).



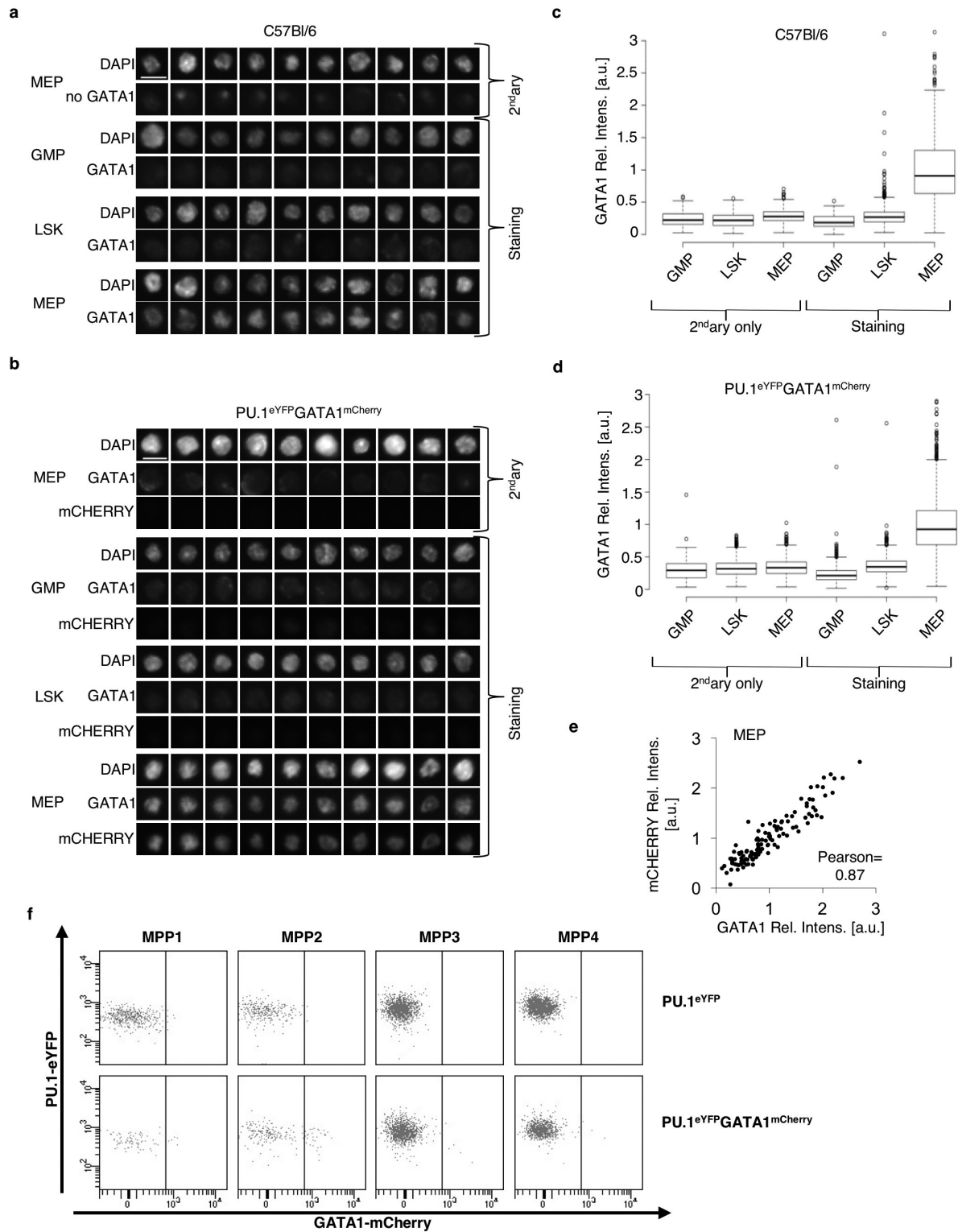
Extended Data Figure 3 | Normal reprogramming capacity of PU.1-eYFP and GATA1-mCherry in both wild-type and PU.1^{eYFP}GATA1^{mCherry} cells. **a, b**, LSK CD34⁺Flt3⁺ were sorted, transduced with lentivirus expressing the indicated proteins and plated in methylcellulose under permissive conditions. Cells from C57Bl/6 mice (**a**) (3 independent experiments) and cells from PU.1-eYFP knock-in mice ($n = 2$; except tdTomato-GATA1 $n = 1$) (**b**) were used. **c, d**, PreMegE cells (**c**) or preGM cells (**d**) from both C57Bl/6 wild-type and PU.1^{eYFP}GATA1^{mCherry} (PU/GATA) knock-in mice were sorted and transduced with mock, PU.1 or PU.1-eYFP expressing lentivirus (**c**) or with mock, GATA1 or GATA1-mCherry expressing lentivirus (**d**), respectively. After 24 h, cells were seeded in methylcellulose under permissive conditions. Colonies were scored after 8–10 days of culture

(3 independent experiments; except for C57Bl/6, PU/GATA, C57Bl/6 + PU.1 and PU/GATA + PU.1 $n = 4$). There was no significant difference between PU.1 and PU.1-eYFP overexpression in either C57Bl/6 or PU/GATA cells ($P > 0.23$; Kruskal-Wallis test). In contrast, C57Bl/6 and PU/GATA colonies without PU.1-eYFP overexpression were significantly different ($P < 0.05$ for M, Meg and E) (**c**). There was no significant difference between GATA1 and GATA1-mCherry overexpression in either C57Bl/6 or PU/GATA cells ($P > 0.77$; Kruskal-Wallis test). In contrast, C57Bl/6 and PU/GATA colonies without GATA1-mCherry overexpression were significantly different ($P < 0.007$, MegE) (**d**). Data are mean values. GEMMeg, granulocytic, erythroid, monocytic, megakaryocytic; MegE, megakaryocytic-erythroid; Meg, megakaryocytic; E, erythroid; GM, granulocytic-monocytic; M, monocytic; G, granulocytic.



Extended Data Figure 4 | PU.1 is expressed in nuclei of all HSCs, MPPs, GMPs and MEPs of wild-type and PU.1^{eYFP}GATA1^{mCherry} mice, and overlaps with eYFP expression. **a, b**, Indicated cell populations from wild-type C57Bl/6 (**a**) and PU.1^{eYFP}GATA1^{mCherry} (**b**) mice were sorted, fixed and probed with PU.1 antibody followed by staining with secondary antibody. Representative examples from controls ('secondary', without primary antibody) and staining of GMPs, MPPs, HSCs, and MEPs (CD150⁺ progenitors) are shown. DAPI stains nuclei. Scale bar, 10 μ m. **c, d**, Quantifications of relative PU.1 expression levels determined by pixel intensities (rel. intens.). Data includes samples from three independent experiments (biological replicates), each of which was normalized to the mean expression levels of the respective MEP population. Individual data points for **c** are 884 MEP secondary only (475, 148 and 261 data points from the individual experiments), 1,218 MPP secondary only (553, 260

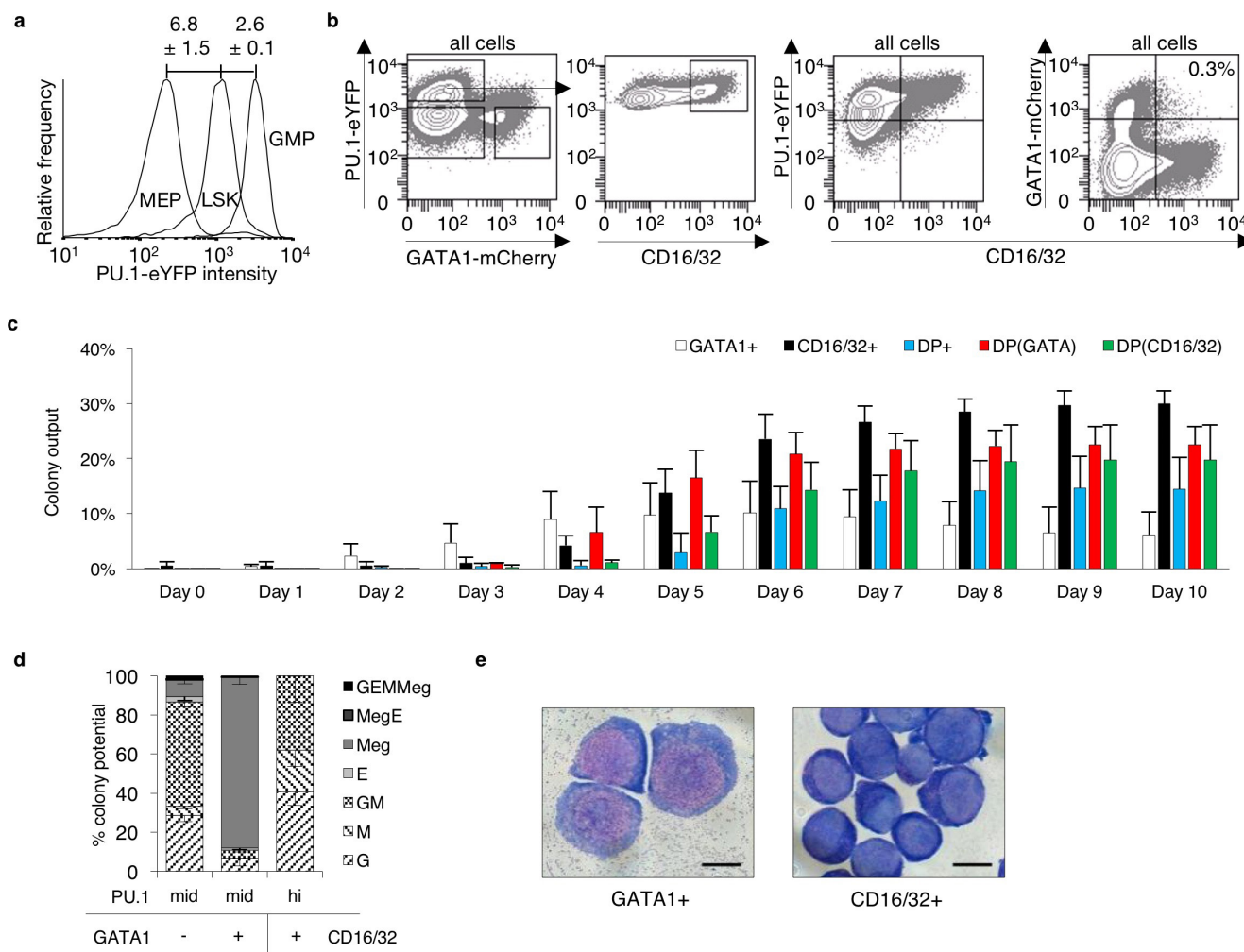
and 405), 755 GMP secondary only (599, 122 and 34), 1,213 MEP (659, 371 and 183), 603 HSCs (360, 194 and 49), 1,458 MPP (749, 283 and 426) and 819 GMP (571, 183 and 65). Individual data for **d** points are 1,530 MEP secondary only (739, 449 and 342 data points from the individual experiments), 1,194 MPP secondary only (547, 394 and 253), 1,866 GMP secondary only (1,616, 126 and 124), 1,521 MEP (518, 581 and 422), 273 HSCs (116, 79 and 78), 1,531 MPP (616, 463 and 452) and 2,339 GMP (1,351, 673 and 315). **e**, Correlation plot of PU.1 and eYFP staining in the indicated cell populations from one experiment. Pixel intensities were normalized to the mean expression in MEPs. Pearson correlation coefficients are displayed. Mean Pearson correlation values (\pm s.d.) of all three independent experiments were 0.82 ± 0.04 (MEP), 0.69 ± 0.10 (HSC), 0.71 ± 0.06 (MPP) and 0.63 ± 0.08 (GMP).



Extended Data Figure 5 | See next page for caption.

Extended Data Figure 5 | GATA1 is expressed in nuclei of MegE committed cells of wild-type and PU.1^{eYFP}GATA1^{mCherry} mice, and overlaps with mCherry expression. **a, b**, Indicated cell populations from wild-type C57Bl/6 (**a**) and PU.1^{eYFP}GATA1^{mCherry} (**b**) mice were sorted, fixed and probed with GATA1 (**a, b**) and mCherry (**b**) antibody followed by staining with secondary antibodies. Representative examples from controls ('secondary', without primary antibody) and staining of GMP, LSK and MEP (CD150⁺ progenitors) are shown. DAPI stains nuclei. Scale bar, 10 μ m. **c, d**, Quantifications of relative GATA1 expression levels determined by pixel intensities. Data includes samples from three independent experiments (biological replicates), each of which was normalized to the mean expression levels of the respective MEP population. Individual data points for **c** are 292 GMP secondary only (56, 188 and 48 data points from the individual experiments), 698 LSK secondary only (155, 287 and 256), 563 MEP secondary only (308, 216

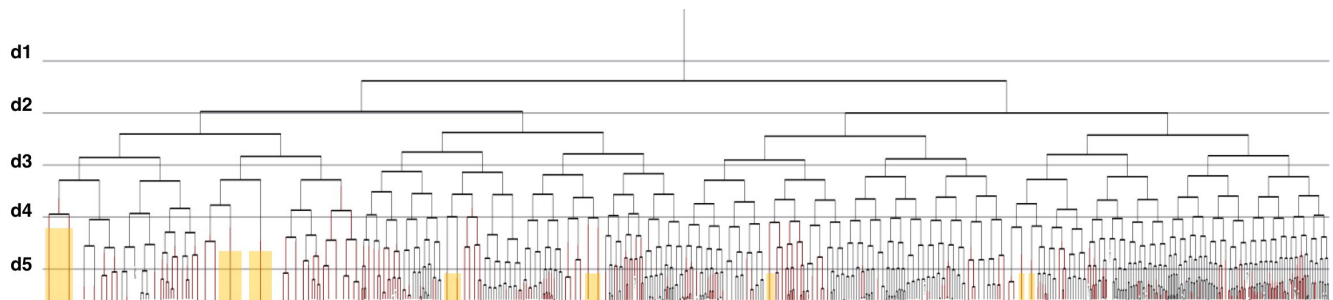
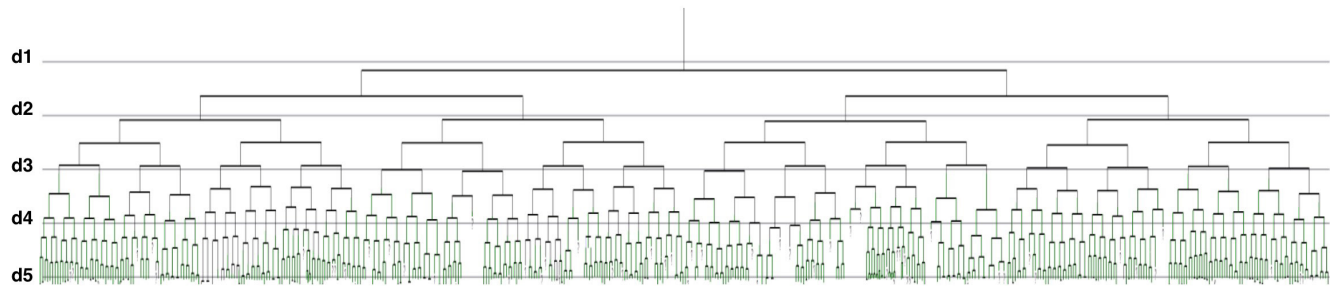
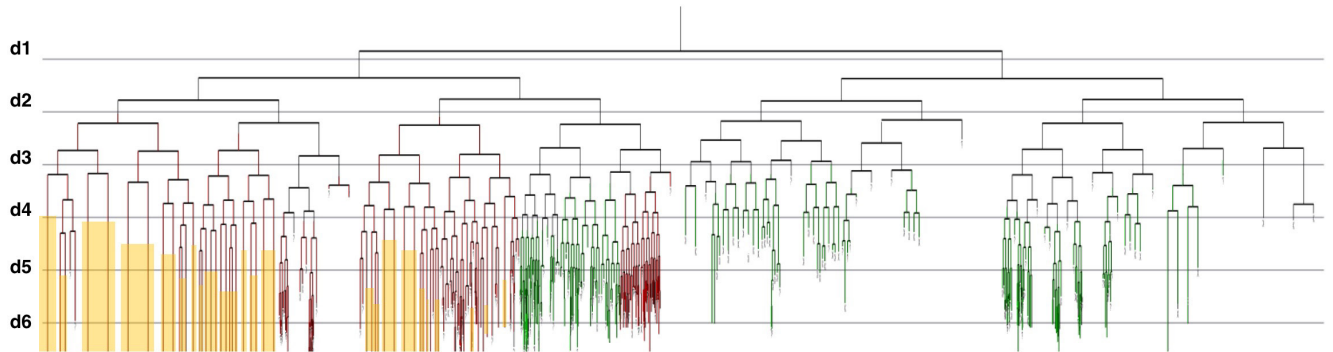
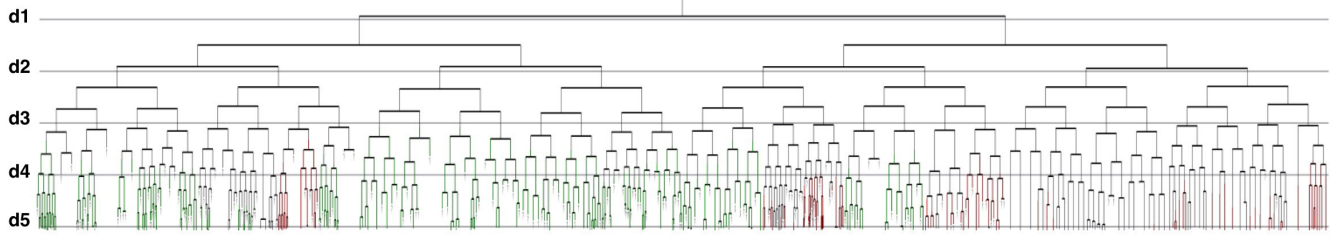
and 39), 344 GMP (64, 158 and 122), 1167 LSK (394, 294 and 479) and 590 MEP (252, 155 and 183). Individual data points for **d** are 485 GMP secondary only (171, 173 and 141 data points from the individual experiments), 1,295 LSK secondary only (360, 552 and 383), 886 MEP secondary only (561, 203 and 122), 462 GMP (73, 114 and 275), 1,184 LSK (252, 441 and 491) and 865 MEP (632, 115 and 118). **e**, Correlation plot of GATA1 and mCherry staining in the indicated cell populations from one experiment. Pixel intensities were normalized to the mean expression in MEPs. The mean Pearson correlation value of all three independent experiments was 0.87 ± 0.06 (s.d.). **f**, Representative example of GATA1–mCherry expression in MPP1–4 from three independent experiments (biological replicates). GATA1–mCherry is also expressed in some MPP2 cells³⁶. These are already committed to the MegE lineage (data not shown). Upper panels PU.1^{eYFP} mouse (negative control), lower panels PU.1^{eYFP}GATA1^{mCherry} mouse.



Extended Data Figure 6 | CD16/32 and GATA1-mCherry expression can be used as a lineage-marker for the GM and MegE lineage, respectively. **a**, Overlapping PU.1-eYFP expression histograms of MEPs, LSKs and GMPs show that PU.1-eYFP levels alone are not sufficient to attribute individual cells to a specific HSPC population. Mean \pm s.d.; $n = 4$ (biological replicates); one representative example shown. **b**, Flow cytometry of day 4 HSC culture. Representative example from three independent experiments (biological replicates). GATA1-mCherry and CD16/32 are mutually exclusive. **c**, Single HSCs were sorted into single wells of a 384-well plate. Colonies were observed for 10 days by one brightfield image per day. Expression of PU.1-eYFP, GATA1-mCherry, and CD16/32 was qualitatively assessed on each day. Colonies were scored into exclusive GATA1-mCherry⁺ (white bars), exclusive CD16/32⁺, or

GATA1-mCherry⁺ and CD16/32 double positive (DP) colonies (57%). DP colonies were further subdivided into colonies that started to express GATA1-mCherry and CD16/32 on the same day (blue bars), expressed GATA1-mCherry at least one day before CD16/32 (red bars), or expressed CD16/32 before GATA1-mCherry (green bars). Missing percentages to 100% mean that colonies have either not expressed any marker yet or that individual colonies have died during the course of 10 days. All surviving colonies have turned on at least one marker (GATA1-mCherry or CD16/32) by day 9. Mean (\pm s.d.) of three biological replicates ($n = 141, 185, 129$ colonies). **d**, Colony potential of sorted cells from **b**, mean \pm s.d. ($n = 3$). **e**, Cytospin of cells from day 4 cultures, representative example from three independent experiments. Scale bars, 10 μ m.

GATA1-mCherry
Megakaryocyte
CD16/32



Extended Data Figure 7 | Additional HSC genealogy examples. Example trees with mixed GM/MegE (1st and 2nd tree), only GM (3rd tree) and only MegE differentiation (4th tree). Compare Fig. 4.

Extended Data Table 1 | PU.1^{eYFP}GATA1^{mCherry} mice are born at normal Mendelian ratios

PU.1 and GATA1	Offspring	Frequency	Expected
PU.1 ^{WT/WT} GATA1 ^{WT/Y}	3	5.8%	6.3%
PU.1 ^{WT/eYFP} GATA1 ^{WT/Y}	5	9.6%	12.5%
PU.1 ^{eYFP/eYFP} GATA1 ^{WT/Y}	1	1.9%	6.3%
PU.1 ^{WT/WT} GATA1 ^{mCherry/Y}	3	5.8%	6.3%
PU.1 ^{WT/eYFP} GATA1 ^{mCherry/Y}	9	17.3%	12.5%
PU.1 ^{eYFP/eYFP} GATA1 ^{mCherry/Y}	3	5.8%	6.3%
PU.1 ^{WT/WT} GATA1 ^{mCherry/WT}	2	3.8%	6.3%
PU.1 ^{eYFP/WT} GATA1 ^{mCherry/WT}	11	21.2%	12.5%
PU.1 ^{eYFP/eYFP} GATA1 ^{mCherry/WT}	5	9.6%	6.3%
PU.1 ^{WT/WT} GATA1 ^{mCherry/mCherry}	2	3.8%	6.3%
PU.1 ^{eYFP/WT} GATA1 ^{mCherry/mCherry}	5	9.6%	12.5%
PU.1 ^{eYFP/eYFP} GATA1 ^{mCherry/mCherry}	3	5.8%	6.3%
PU.1			
PU.1 ^{WT/WT}	10	19.2%	25.0%
PU.1 ^{eYFP/WT}	30	57.7%	50.0%
PU.1 ^{eYFP/eYFP}	12	23.1%	25.0%
GATA1			
GATA1 ^{WT/Y}	9	17.3%	25.0%
GATA1 ^{mCherry/Y}	15	28.8%	25.0%
GATA1 ^{mCherry/WT}	18	34.6%	25.0%
GATA1 ^{mCherry/mCherry}	10	19.2%	25.0%

Offspring of the mating male PU.1^{eYFP/WT} GATA1^{mCherry/Y} mice with female PU.1^{eYFP/WT} GATA1^{mCherry/WT} mice. Shown are real and expected frequencies for the respective genotypes.

Extended Data Table 2 | Protein abundance in different cell populations in flow cytometry and imaging

Flow Cytometry		
Cell population	PU.1eYFP	GATA1mCHERRY
PU.1 ⁺ GATA1 ⁻ LK progenitor	43.1 ± 10.6 × 10 ³	NA
E14.5 GATA1 ⁺ fetal liver	NA	23.0 ± 9.8 × 10 ³
PU.1 ^{high} GATA1 ⁻	47.3 ± 12.9 × 10 ³	-- / < negative gate
PU.1 ^{mid} GATA1 ^{mid}	5.9 ± 1.8 × 10 ³	25.5 ± 12.3 × 10 ³
PU.1 ⁻ GATA1 ^{high}	-- / < negative gate	54.6 ± 23.8 × 10 ³
HSC	8.1 ± 2.1 × 10 ³	-- / < negative gate
LSK	16.4 ± 4.3 × 10 ³	-- / < negative gate
GMP	42.7 ± 11.7 × 10 ³	-- / < negative gate
MEP	-- / < negative gate	49.4 ± 21.4 × 10 ³
Negative gate (GMP)	4.4 ± 1.2 × 10 ³	6.5 ± 3.8 × 10 ³
Negative gate (MEP)	4.7 ± 1.5 × 10 ³	8.4 ± 4.6 × 10 ³

Imaging		
Cell population	PU.1eYFP	GATA1mCHERRY
PU.1 ^{high} GATA1 ⁻	47.3 ± 12.9 × 10 ³	-- / < negative gate
PU.1 ^{mid} GATA1 ^{mid}	5.5 ± 4.4 × 10 ³	25.6 ± 16.3 × 10 ³
PU.1 ⁻ GATA1 ^{high}	-- / < negative gate	54.6 ± 23.8 × 10 ³
HSC	8.1 ± 2.1 × 10 ³	-- / < negative gate
GMP	40.1 ± 4.7 × 10 ³	-- / < negative gate
Negative gate (in silico)	1.1 ± 2.0 × 10 ³	1.9 ± 4.4 × 10 ³

Calculated PU.1-eYFP and GATA1-mCherry protein molecule numbers both, for flow cytometry and imaging for the respective cell populations, as well as the negative gates (that is, the detection thresholds).

Extended Data Table 3 | Data overview from time-lapse movies

	Movie 1	Movie 2	Movie 3	Movie 4	All Movies
Starting cells	63	61	62	70	256
Early Apoptosis (<48h)	15	16	8	7	46
Early Apoptosis (<48h) %	23,8%	26,2%	12,9%	10,0%	18,2%
Lost without onset	5	1	0	3	9
Lost without onset %	7,9%	1,6%	0,0%	4,3%	3,5%
Trees with onsets	34	32	31	51	148
Trees with onsets %	54,0%	52,5%	50,0%	72,9%	57,4%
Trees without onsets	9	12	23	9	53
Trees without onsets %	14,3%	19,7%	37,1%	12,9%	21,0%
GM onsets	227	146	163	544	1080
ME onsets	89	230	93	269	681
Starting cells with >0 GM onset	31	20	19	31	101
Starting cells with >0 ME onset	8	16	17	29	70
Double positive trees	6	4	5	8	23
Trees with >1 GM (and no ME) onset	25	16	14	23	78
Trees with >1 ME (and no GM) onset	2	12	12	21	47
Deepest tracked division per tree (Mean)	6,6	6,0	7,0	7,8	6,9
Standard Deviation	4,5	4,4	3,2	3,7	4,0
Trees with max. 0 divisions	12	11	3	4	30
Trees with max. 1 divisions	2	3	1	3	9
Trees with max. 2 divisions	0	1	3	1	5
Trees with max. 3 divisions	0	0	1	2	3
Trees with max. 4 divisions	1	2	4	0	7
Trees with max. 5 divisions	1	1	4	4	10
Trees with max. 6 divisions	1	0	6	6	13
Trees with max. 7 divisions	2	2	4	4	12
Trees with max. 8 divisions	6	3	4	3	16
Trees with max. 9 divisions	8	8	12	6	34
Trees with max. 10 divisions	5	9	7	14	35
Trees with max. 11 divisions	6	2	3	5	16
Trees with max. 12 divisions	3	2	2	7	14
Trees with max. 13 divisions	1	0	0	1	2
Mean Onset per Tree (if >0 onsets)	6,0	4,1	5,1	5,4	5,2
Mean Standard Deviation	2,6	3,5	2,8	3,1	3,0
Mean Onset Generation 0	4	10	4	10	28
Mean Onset Generation 1	0	3	2	0	5
Mean Onset Generation 2	0	1	1	0	2
Mean Onset Generation 3	1	0	1	1	3
Mean Onset Generation 4	0	2	2	2	6
Mean Onset Generation 5	4	1	4	7	16
Mean Onset Generation 6	9	2	4	8	23
Mean Onset Generation 7	7	6	7	8	28
Mean Onset Generation 8	5	5	5	9	24
Mean Onset Generation 9	3	2	1	4	10
Mean Onset Generation 10	1	0	0	1	2
Mean Onset Generation 11	0	0	0	0	0
Mean Onset Generation 12	0	0	0	0	0
Mean Onset Generation 13	0	0	0	0	0

Information on tracked colonies per independent movie and their fate outcome. Number of colonies regarding their latest tracked division and mean marker onset (GATA1-mCherry and/or CD16/32). General tree fates include 'apoptosis' (<48 h after movie start), 'lost' (no information because cell identity lost), 'onset' (at least one marker onset per tracked tree) and 'no onset' (no marker onset in tracked tree).

TTC39B deficiency stabilizes LXR reducing both atherosclerosis and steatohepatitis

Joanne Hsieh^{1*}, Masahiro Koseki^{1,2*}, Matthew M. Molusky¹, Emi Yakushiji¹, Ikuyo Ichi³, Marit Westerterp¹, Jahangir Iqbal⁴, Robin B. Chan⁵, Sandra Abramowicz¹, Liana Tascu¹, Shunichi Takiguchi^{6,7}, Shizuya Yamashita², Carrie L. Welch¹, Gilbert Di Paolo⁵, M. Mahmood Hussain⁴, Jay H. Lefkowitz⁵, Daniel J. Rader^{6,7} & Alan R. Tall¹

Cellular mechanisms that mediate steatohepatitis, an increasingly prevalent condition in the Western world for which no therapies are available¹, are poorly understood. Despite the fact that its synthetic agonists induce fatty liver, the liver X receptor (LXR) transcription factor remains a target of interest because of its anti-atherogenic, cholesterol removal, and anti-inflammatory activities. Here we show that tetratricopeptide repeat domain protein 39B (*Ttc39b*, C9orf52) (*T39*), a high-density lipoprotein gene discovered in human genome-wide association studies², promotes the ubiquitination and degradation of LXR. Chow-fed mice lacking *T39* (*T39*^{-/-}) display increased high-density lipoprotein cholesterol levels associated with increased enterocyte ATP-binding cassette transporter A1 (*Abca1*) expression and increased LXR protein without change in LXR messenger RNA. When challenged with a high fat/high cholesterol/bile salt diet, *T39*^{-/-} mice or mice with hepatocyte-specific *T39* deficiency show increased hepatic LXR protein and target gene expression, and unexpectedly protection from steatohepatitis and death. Mice fed a Western-type diet and lacking low-density lipoprotein receptor (*Ldlr*^{-/-} *T39*^{-/-}) show decreased fatty liver, increased high-density lipoprotein, decreased low-density lipoprotein, and reduced atherosclerosis. In addition to increasing hepatic *Abcg5/8* expression and limiting dietary cholesterol absorption, *T39* deficiency inhibits hepatic sterol regulatory element-binding protein 1 (SREBP-1, ADD1) processing. This is explained by an increase in microsomal phospholipids containing polyunsaturated fatty acids, linked to an LXR α -dependent increase in expression of enzymes mediating phosphatidylcholine biosynthesis and incorporation of polyunsaturated fatty acids into phospholipids. The preservation of endogenous LXR protein activates a beneficial profile of gene expression that promotes cholesterol removal and inhibits lipogenesis. *T39* inhibition could be an effective strategy for reducing both steatohepatitis and atherosclerosis.

Genome-wide association studies have uncovered a plethora of novel genetic loci associated with alterations in plasma lipoprotein levels^{2,3} that have potential to provide insights into metabolic diseases such as atherosclerosis and fatty liver. Single nucleotide polymorphisms in intron 1 of *T39* were associated with reduced hepatic *T39* messenger RNA (mRNA) and increased high-density lipoprotein (HDL) cholesterol levels². However, the only clue to the cellular functions of *T39* is that it contains three consecutive tetratricopeptide repeat motifs, suggesting it might function as a scaffolding protein mediating the association of HDL-regulating proteins. *T39* mRNA was highly expressed in liver and small intestine of chow-fed wild-type (WT) mice and was reduced by >90% in *T39*^{-/-} mice (Extended Data Fig. 1). HDL cholesterol levels were increased by ~22% in chow-fed *T39*^{-/-} mice compared with WT (Extended Data Fig. 2a), while non-HDL cholesterol

and triglyceride levels were unchanged (data not shown). *T39*^{-/-} mice challenged with 3 weeks of the high fat/high cholesterol/bile salt (HF/HC/BS) diet had a 42% increase in HDL cholesterol levels (Extended Data Fig. 2a), a 45% increase in apolipoprotein A-1 (ApoA-1), the major protein component of HDL particles (Extended Data Fig. 2b), decreased very low-density lipoprotein (VLDL)/chylomicron cholesterol levels (Extended Data Fig. 2c), and no difference in plasma triglyceride levels (data not shown). Gene expression microarrays of the liver of chow-fed *T39*^{-/-} and WT mice showed no significant differences in genes potentially involved in the regulation of HDL, including *Apoa1*, scavenger receptor b1, hepatic lipase, and *Abca1*, which accounts for over 90% of HDL formation⁴ (data not shown).

There is longstanding evidence that the intestine makes a substantial contribution to the production of HDL^{5,6}. Small intestinal enterocytes from chow-fed *T39*^{-/-} mice showed increased *Abca1* mRNA and protein (Extended Data Fig. 2d, e and Fig. 1a). Protein levels of both isoforms of LXR, the major transcriptional activator of *Abca1* (ref. 7), were dramatically increased (Fig. 1a and Extended Data Fig. 2e), while *Lxr α* (*Nr1h3*, *RLD1*) and *Lxr β* (*Nr1h2*, *UR*) mRNA levels were unchanged (Extended Data Fig. 2d). We also observed induction of other intestinal LXR target genes including inducible degrader of LDLR (*Idol*, *Myli1*)⁸ (Extended Data Fig. 2d). There was increased incorporation of [³H]cholesterol into HDL-sized particles secreted from *T39*^{-/-} enterocytes isolated from chow- (Extended Data Fig. 2f) and HF/HC/BS-diet-fed mice (Extended Data Fig. 2g). On the chow diet, enterocyte-specific *T39* deletion in *Villin-Cre*⁺ *T39*^{fl/fl} mice raised HDL cholesterol, whereas hepatocyte-specific *T39* deletion in *Albumin-Cre*⁺ *T39*^{fl/fl} mice had no effect, confirming the intestinal contribution to increased HDL (Fig. 1b). *T39* deficiency did not yield any difference in HDL cholesterol on the *Lxr α* ^{-/-} background (data not shown). Together, these findings suggest that the major mechanism responsible for increased HDL levels in chow-fed *T39*^{-/-} mice is increased intestinal expression of *Abca1*, secondary to a post-transcriptional induction of LXR protein. Hepatocyte *T39* deficiency did, however, increase HDL cholesterol in mice fed the HF/HC/BS diet (Fig. 1b), which was reversed by LXR α deficiency (data not shown). This suggests that under inflammatory conditions induced by the HF/HC/BS diet, the liver also contributes to the HDL phenotype of *T39*^{-/-} mice, consistent with the previous report that *T39* knockdown mediated by adenovirus, which targets the liver and is inflammatory, raised HDL².

HF/HC/BS diets have been used as a model of steatohepatitis resembling human non-alcoholic steatohepatitis⁹. After 20 weeks of HF/HC/BS feeding, we noticed a fourfold reduction in mortality among *T39*^{-/-} mice ($P < 0.05$) (Fig. 1c), accompanied by decreased circulating alanine aminotransferase (ALT) levels (Extended Data Fig. 3a). Livers were smaller and less pale in the *T39*^{-/-} mice than controls (Extended

¹Division of Molecular Medicine, Department of Medicine, Columbia University, New York, New York 10032, USA. ²Cardiovascular Medicine, Osaka University Graduate School of Medicine, Suita, Osaka 565-0871, Japan. ³Faculty of Core Research, Ochanomizu University, Bunkyo-ku, Tokyo 112-8610, Japan. ⁴Department of Cell Biology, State University of New York Health Science Center at Brooklyn (SUNY Downstate Medical Center), Brooklyn, New York 11203, USA. ⁵Department of Pathology and Cell Biology, Columbia University, New York, New York 10032, USA. ⁶Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia 19104, USA. ⁷Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia 19104, USA.

*These authors contributed equally to this work.

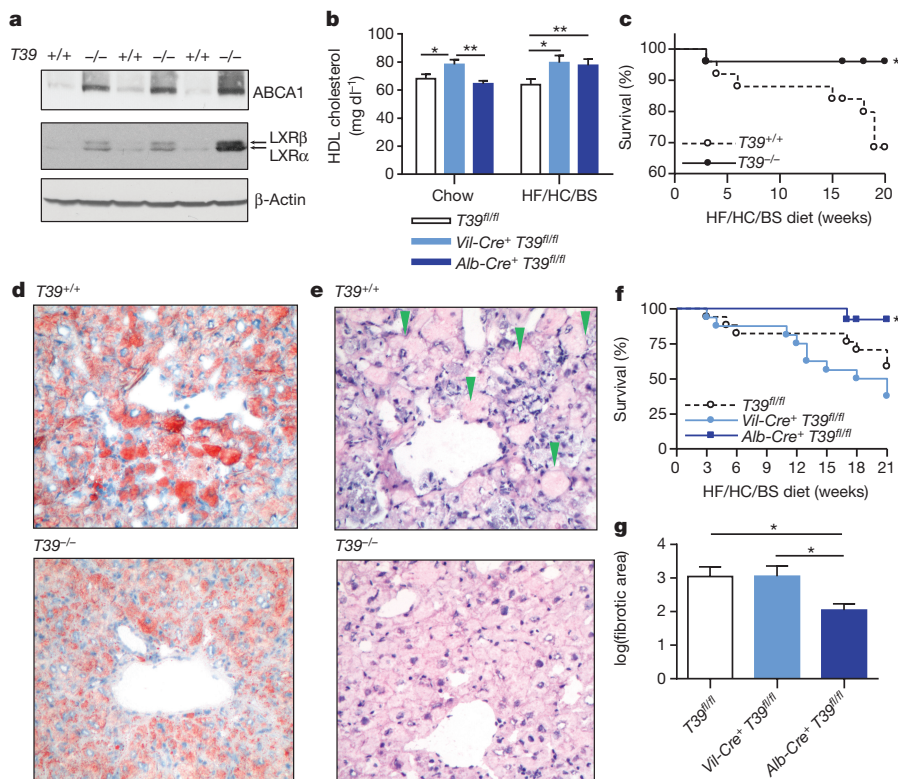


Figure 1 | Increased HDL cholesterol and protection from steatohepatitis in T39-deficient mice. **a**, Enterocyte protein expression of ABCA1 (top), both LXR isoforms (middle), and β -actin (bottom). **b**, HDL-cholesterol of male tissue-specific T39 knockout mice fed chow or HF/HC/BS diet for 2 weeks. **c**, Mortality of WT and whole-body T39 knockout mice fed the HF/HC/BS diet, $*P < 0.05$ based on log rank test of the Kaplan–Meier curve, $n = 25$ per genotype. **d**, Representative Oil Red O and **(e)** haematoxylin and eosin stains of hepatic sections from female mice fed HF/HC/BS diet for 18 weeks. Green arrowheads point to ballooned hepatocytes. **f**, Mortality of tissue-specific T39 knockout on the HF/HC/BS diet. $*P < 0.05$ based on log rank test of the Kaplan–Meier curve, $n = 13$ –17 per genotype. For **b** and **f**, $n = 17$ $T39^{fl/fl}$, 16 $Villin-Cre^{+}T39^{fl/fl}$, and 13 $Albumin-Cre^{+}T39^{fl/fl}$. **g**, Hepatic fibrotic area based on Masson's trichrome staining, $n = 8$ per genotype. Data are represented as mean \pm s.e.m., unless indicated otherwise, $*P < 0.05$ and $**P < 0.01$ by one-way analysis of variance (ANOVA). See Supplementary Fig. 1 for gel source data.

Data Fig. 3b), while there were no differences in body weight or gonadal fat pad weight (data not shown). The livers of $T39^{-/-}$ mice had less Oil Red O staining (Fig. 1d), reflecting diminished hepatic triglyceride (Extended Data Fig. 3c) and cholesteryl ester (Extended Data Fig. 3d) accumulation, fewer inflammatory foci consisting of neutrophils and lymphocytes (Fig. 1e and Extended Data Fig. 3e), less hepatocellular ballooning degeneration (Fig. 1e and Extended Data Fig. 3f), and less hepatocyte proliferation in $T39^{-/-}$ mice (Extended Data Fig. 3g). Mortality studies in tissue-specific T39 knockout mice revealed that protection was entirely due to hepatic T39 deficiency

(Fig. 1f). The livers of $Albumin-Cre^{+}T39^{fl/fl}$ mice had less perisinusoidal and periportal fibrosis than the $T39^{fl/fl}$ and $Villin-Cre^{+}T39^{fl/fl}$ mice (Fig. 1g and Extended Data Fig. 4a, b), reduced inflammation (Extended Data Fig. 4a, c), and less hepatocellular ballooning (Extended Data Fig. 4a, d). In the livers of $T39^{-/-}$ mice fed HC/HF/BS for 6 weeks, many LXR targets were upregulated, including *Abcg5/8*, stearoyl coenzyme A desaturase 1 (*Scd1*), elongation of very long chain fatty acids protein 5 (*Elovl5*), insulin induced gene 2a (*Insig2a*), and lysophatidylcholine acyltransferase 3 (*Lpcat3*, *Mboat5*)¹⁰, and these increases were reversed in $Lxra^{-/-}T39^{-/-}$ mice (Fig. 2a). However,

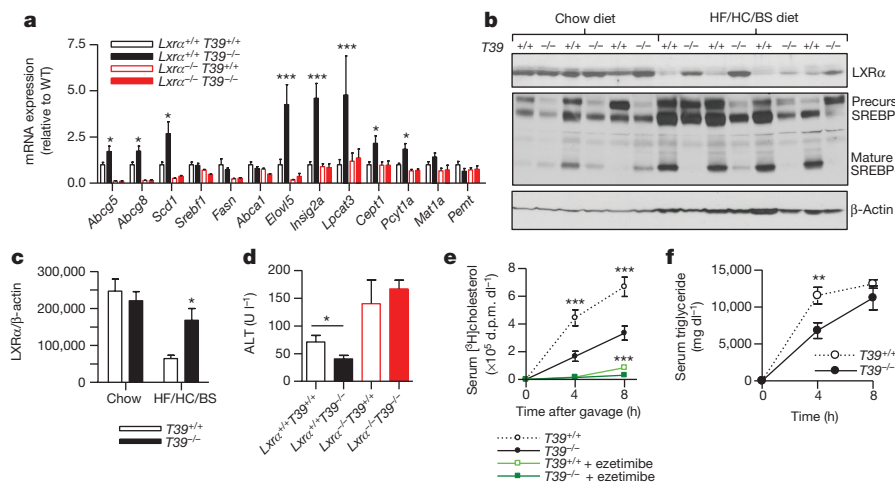


Figure 2 | Post-transcriptional LXR activation and decreased dietary cholesterol absorption in T39-deficient mice. **a**, Hepatic gene expression in fasted WT and whole-body T39 knockout mice with/without LXRα fed HF/HC/BS diet for 6 weeks, $***P < 0.001$ by two-way ANOVA. **b**, Hepatic protein expression of WT or whole-body T39 knockout mice fed HF/HC/BS diet for 5 weeks, showing LXRα (top), both the microsomal and nuclear forms of SREBP-1 (middle), and β -actin (bottom). **c**, Quantification of LXRα protein, $n = 4$ per genotype per diet. **d**, Serum ALT levels of WT and whole-body T39 knockout mice with/without LXRα fed HF/HC/BS

diet for 2 weeks. For **a** and **c**, $n = 7$ $Lxra^{+/+}T39^{+/+}$, 11 $Lxra^{+/+}T39^{-/-}$, 8 $Lxra^{-/-}T39^{+/+}$, and 4 $Lxra^{-/-}T39^{-/-}$. **e**, Short-term dietary cholesterol absorption in WT and whole-body T39 knockout mice fed HF/HC/BS diet for 4 weeks, with/without ezetimibe treatment, $n = 5$ per genotype, $***P < 0.001$ by two-way ANOVA; d.p.m., disintegrations per minute. **f**, Postprandial triglyceride secretion in WT and whole-body T39 knockout mice, $n = 5$ per genotype, $**P < 0.01$ by two-way ANOVA. Data are represented as mean \pm s.e.m., unless indicated otherwise, $*P < 0.05$ and $**P < 0.01$ by two-tailed t -test. See Supplementary Fig. 1 for gel source data.

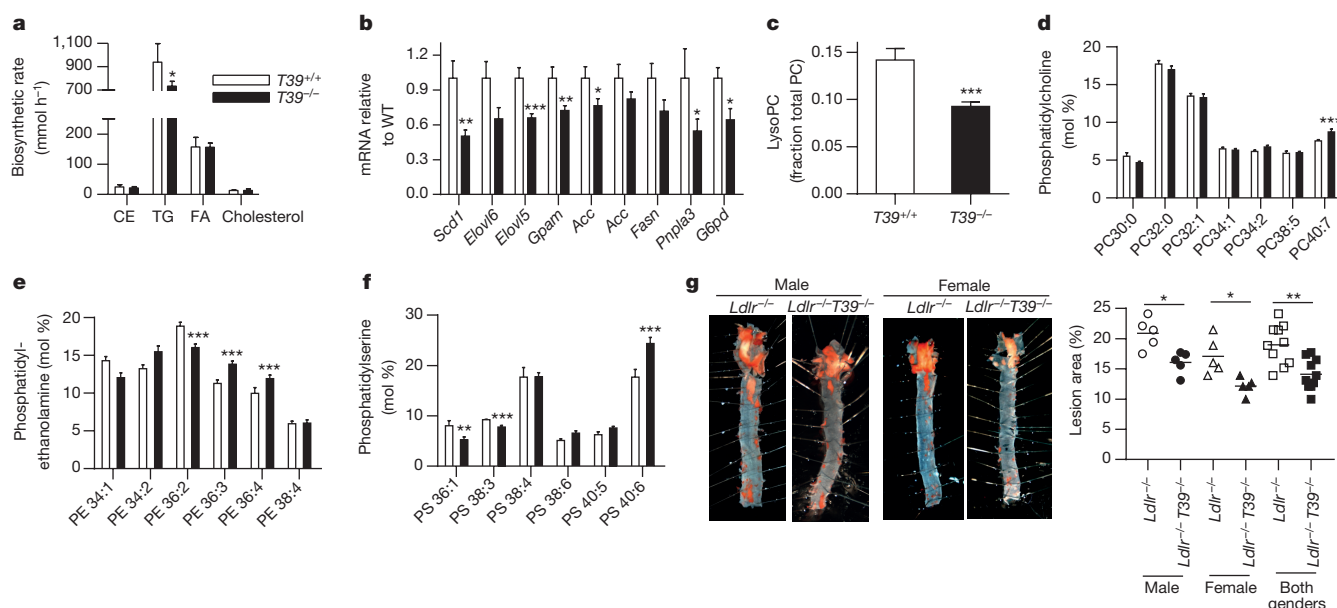


Figure 3 | Changes in phosphatidylcholine metabolism decrease lipogenesis along with reduced atherosclerosis. **a**, Hepatic lipid biosynthetic rate in WT or whole-body T39 knockout mice fed WTD for 16 weeks, $n = 6$ WT and 8 $T39^{-/-}$, $^{**}P < 0.01$ by two-ANOVA. **b**, Hepatic lipogenic gene expression in fasted-refed mice fed WTD for 16 weeks, $n = 7$ WT and 13 $T39^{-/-}$. **c**, Lysophosphatidylcholine:phosphatidylcholine ratio and the fatty-acid composition of **(d)** phosphatidylcholine (PC),

expression of lipogenic genes *Srebf1* and fatty acid synthase (*Fasn*) was unchanged (Fig. 2a). There was no difference in LXR α protein expression between WT and $T39^{-/-}$ mice fed chow (Fig. 2b). However, whereas feeding a HF/HC/BS diet reduced hepatic LXR α protein levels in WT mice, LXR protein levels in $T39^{-/-}$ livers were largely preserved (Fig. 2b, c), without any difference in *Lxr α* mRNA level (Extended Data Fig. 5a). Similar mRNA levels of sulfotransferase 2b1 (*Sult2b1*) and *Abcc1* were found in livers from both genotypes, suggesting that decreased LXR ligand sulfation and export¹¹ did not account for increased LXR target gene expression in $T39^{-/-}$ mice (data not shown). Also, hepatic levels of natural LXR ligands, including 24S-, 25-, 27-hydroxycholesterol and desmosterol, were decreased in whole-body (Extended Data Fig. 5b) and unchanged in hepatocyte-specific T39 knockout mice (Extended Data Fig. 5c). Thus, increased hepatic expression of LXR target genes in $T39^{-/-}$ mice (Fig. 2a) was primarily due to a post-transcriptional increase in LXR protein. Differences in serum ALT levels were abrogated in mice lacking LXR α (Fig. 2d), indicating that LXR activation is critical to the hepatoprotective effects of T39 deficiency. $T39^{-/-}$ mice had decreased dietary cholesterol absorption (Fig. 2e), consistent with observations that hepatic and intestinal LXR α activation have been shown to decrease dietary cholesterol absorption^{12,13}. There was also delayed postprandial plasma triglyceride accumulation in $T39^{-/-}$ mice (Fig. 2f). Reduced cholesterol absorption was LXR α -dependent (Extended Data Fig. 5d), and associated with reduced accumulation of exogenously derived hepatotoxic oxysterols such as 7 β -hydroxycholesterol, and 7-ketocholesterol (Extended Data Fig. 5b). The decrease in cholesterol absorption was not associated with a change in intestinal Niemann–Pick C1 like 1 (*Npc1l1*) expression (Extended Data Fig. 5e) and persisted in the presence of the NPC1L1 inhibitor ezetimibe (Fig. 2e), suggesting that these changes are attributable to upregulation of the sterol exporters ABCG5/8 (Fig. 2a), which reduces fractional cholesterol absorption¹⁴. Consistent with ABCG5/8 induction¹⁴, hepatic concentrations of the plant sterol campesterol were lower in $T39^{-/-}$ mice (Extended Data Fig. 5f), and there was enhanced reverse cholesterol transport (Extended Data Fig. 5g). While the decrease in sterol absorption could contribute to

(e) phosphatidylethanolamine (PE), and **(f)** phosphatidylserine (PS) in microsomes of livers from mice fed HF/HC/BS diet for 6 weeks, $n = 7$ WT and 10 $T39^{-/-}$. **g**, Atherosclerotic lesions of WT and whole-body T39 knockout mice lacking LDLR fed WTD for 20 weeks, $n = 5$ per genotype per gender. Data are represented as mean \pm s.e.m., unless indicated otherwise, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ by two-tailed *t*-test. See Supplementary Fig. 1 for gel source data.

liver protection, both *Albumin-Cre*⁺ $T39^{fl/fl}$ and *Villin-Cre*⁺ $T39^{fl/fl}$ mice exhibited significantly decreased cholesterol absorption (Extended Data Fig. 5h), whereas only the liver knockout was protective (Fig. 1f), suggesting there must be another mechanism contributing to the hepatoprotective effect of T39 deficiency.

LXR activation by synthetic agonists has traditionally been associated with hypertriglyceridemia and hepatic steatosis, owing to induction of SREBP-1c, the transcription factor that controls lipogenic gene expression¹⁵. However, in enterocytes and livers of $T39^{-/-}$ mice, *Srebf1* mRNA and precursor protein were not induced (Extended Data Fig. 2d and Fig. 2a, b). Notably, there was a dramatic inhibition of SREBP-1 processing to the smaller, transcriptionally active form in the livers of $T39^{-/-}$ mice fed the HF/HC/BS diet (Fig. 2b). The processing of SREBP-2, the transcription factor that regulates cholesterol synthesis, was not affected (Extended Data Fig. 6a). The decrease in the mature form of SREBP-1 associated with T39 deficiency was LXR α dependent (Extended Data Fig. 6b). $T39^{-/-}$ mice fed the Western-type diet (WTD) also had reduced hepatic triglyceride (Extended Data Fig. 6c) and cholesterol (Extended Data Fig. 6d) accumulation, and inhibition of SREBP-1 processing (Extended Data Fig. 6e). Potentially contributing to decreased SREBP-1 processing, in the fasting state, there was marked LXR α -dependent induction of *Insig2a* (Fig. 2a), the polytopic protein that retains SREBP cleavage-activating protein (SCAP) and therefore SREBP in the endoplasmic reticulum, thereby preventing SREBP proteolytic cleavage. However, *Insig2a* is only induced by LXR in the absence of insulin¹⁶, while SREBP-1 activation is prominent in the postprandial state when insulin levels are elevated, suggesting there must be a separate, more physiologically relevant mechanism regulating SREBP-1 maturation.

In a fasting–refeeding experiment, WTD-fed $T39^{-/-}$ mice displayed significantly decreased hepatic triglyceride synthesis *in vivo* (Fig. 3a). The reduction in triglyceride synthesis was accompanied by decreased expression of many lipogenic genes, including *Scd1*, *Elovl5*, acetyl-CoA carboxylase α (*Acc α*), glycerol-3-phosphate acyltransferase (*Gpat1*, *Gpat*), and glucose-6-phosphate dehydrogenase (*G6pd*), as well as patatin-like phospholipase domain-containing protein 3 (*Pnpla3*, adiponutrin), which is an SREBP-1 target implicated in human hepatic

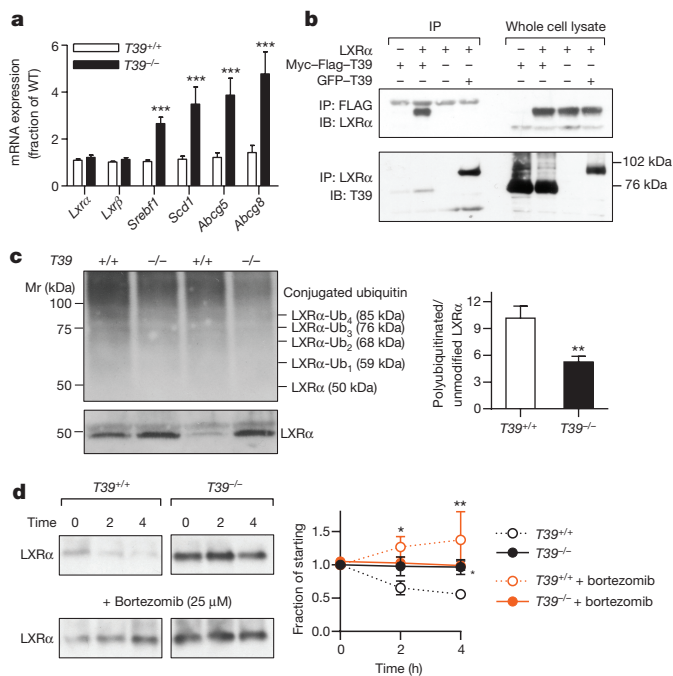


Figure 4 | T39 deficiency stabilizes LXRα in hepatocytes. **a**, Primary WT and T39 knockout hepatocyte mRNA expression, $n = 4$ per genotype. **b**, A protein complex containing LXRα and T39 as demonstrated by T39 immunoprecipitation (top) and the converse LXRα immunoprecipitation (bottom) shown on the left, while whole cell lysates are shown on the right. **c**, *In vivo* ubiquitination of endogenous LXRα in livers of WT or whole-body T39 knockout mice fed HF/HC/BS diet for 6 weeks and treated with a proteasomal inhibitor, ubiquitination signal of GW3965-mediated pulldown (top), unmodified LXRα (bottom), and quantification of the polyubiquitinated LXRα:unmodified LXRα ratio (right), $n = 10$ per genotype. **d**, Myc-Flag-LXRα turnover in WT or T39 knockout hepatocytes with quantification of LXRα shown on the right, * $P < 0.05$ and ** $P < 0.01$ compared with WT by two-way ANOVA. Data are represented as mean \pm s.e.m., unless indicated otherwise, *** $P < 0.01$, *** $P < 0.001$ by two-tailed *t*-test. See Supplementary Fig. 1 for gel source data.

steatosis^{17,18} (Fig. 3b). *Albumin-Cre*⁺*T39*^{fl/fl} mice also had significant reductions in postprandial hepatic lipogenic gene expression (Extended Data Fig. 6f). Since exogenous free fatty acids are the major source of hepatic triglyceride in mice fed a high fat diet¹⁹, the decrease in [³H] triglyceride observed in Fig. 3d is probably due to lower rates of esterification of exogenous FA, possibly because of decreased hepatic *Gpat1* expression (Fig. 3b) and delayed dietary triglyceride absorption (Fig. 2f). Hepatic insulin sensitivity did not differ between *T39*^{-/-} and WT mice, as glucose tolerance (Extended Data Fig. 6g), hepatic gluconeogenesis (Extended Data Fig. 6h), and insulin-induced Akt phosphorylation were similar in WT and *T39*^{-/-} mice (Extended Data Fig. 6i).

Microsomal lipid composition changes could have a sustained effect on SREBP-1 processing in the fed state. Among the most significant changes in the membrane lipids was a 30% increase in phosphatidylcholine content in *T39*^{-/-} mice, with no differences in phosphatidylethanolamine and phosphatidylserine (data not shown). The increased in phosphatidylcholine may be attributable to LXRα-dependent upregulation of the Kennedy pathway enzymes, phosphate cytidyltransferase 1a (*Pcyt1a*) and choline/ethanolamine phosphotransferase 1 (*Cept1*) (Fig. 2a). Enzymes involved in an alternative methylation-dependent phosphatidylcholine synthesis pathway were not changed (Fig. 2a). Phosphatidylcholine synthesis is an important regulator of SREBP-1 proteolytic maturation in the liver²⁰. In the endoplasmic reticulum, LPCAT3-mediated phospholipid remodelling preferentially acylates the *sn*-2 position of lysophosphatidylcholine, lysophosphatidylethanolamine, and lysophosphatidylserine with polyunsaturated fatty acids (PUFAs)²¹. There was a decreased ratio of lysophosphatidylcholine to

phosphatidylcholine (Fig. 3c), and phosphatidylcholine, phosphatidylethanolamine, and phosphatidylserine species all had significantly increased PUFA content (Fig. 3d–f), consistent with *Lpcat3* upregulation²². PUFAs have been demonstrated to inhibit the transcription and processing of SREBP-1c²³. Also, LPCAT3 activity ameliorates endoplasmic reticulum stress¹⁰, which affects SREBP-1 processing²⁴. Together these data suggest that an LXRα-dependent increase in phosphatidylcholine synthesis and incorporation of PUFA into multiple phospholipid species in *T39*^{-/-} mice may be a major factor inhibiting the development of non-alcoholic fatty liver disease.

T39^{-/-} mice were crossed into the *Ldlr*^{-/-} background and fed a WTD to determine whether the anti-atherogenic effects of LXR activation were maintained. Male *Ldlr*^{-/-}*T39*^{-/-} mice had significantly lower total cholesterol and higher HDL cholesterol (Extended Data Fig. 7a). Male and female *Ldlr*^{-/-}*T39*^{-/-} mice both showed a reduction in LDL cholesterol (Extended Data Fig. 7a, b), but plasma triglyceride levels were not different (data not shown). Atherosclerotic lesion area evaluated by *en face* Oil Red O staining of the aorta was significantly reduced in *Ldlr*^{-/-}*T39*^{-/-} mice compared with *Ldlr*^{-/-} controls (Fig. 3g). Although there was no difference in overall lesion area in the proximal aorta (Extended Data Fig. 7c), atherosclerosis was less advanced as judged by reduced lesion complexity in *Ldlr*^{-/-}*T39*^{-/-} mice (Extended Data Fig. 7d). Thus, T39 deficiency confers protection from atherosclerosis and fatty liver following challenge with the WTD.

In contrast to the liver, in cultured primary hepatocytes from chow-fed *T39*^{-/-} mice, the mRNA levels of LXR target genes (*Srebf1*, *Scd1*, *Abcg5*, and *Abcg8*) were increased, while expression of *Lxrα* and *Lxrβ* were unchanged (Fig. 4a). LXRα protein was increased in the cytosol and nucleus of *T39*^{-/-} hepatocytes (Extended Data Fig. 8a). We used the obligate heterodimer partner of LXR, retinoid X receptor (RXR) as a surrogate for the presence of LXR over LXR response elements (LXREs). There seemed to be increased occupancy by LXR/RXR over LXREs of several LXR targets both in the basal state and after GW3965 treatment in *T39*^{-/-} compared with WT hepatocytes (Extended Data Fig. 8b). In co-overexpression studies, LXRα co-immunoprecipitated with T39 regardless of the epitope tag size (Fig. 4b), indicating that LXRα and T39 coexist in the same protein complex. We discovered that endogenous LXR could be isolated using immobilized GW3965 (Extended Data Fig. 8c). In mice treated with a proteasome inhibitor, there was significantly less polyubiquitinated LXRα relative to unmodified LXRα in *T39*^{-/-} livers (Fig. 4c). There was a significantly slower rate of LXRα turnover in *T39*^{-/-} than WT hepatocytes (Fig. 4d). While the proteasome inhibitor bortezomib increased LXR in WT hepatocytes, it had no effect in *T39*^{-/-} hepatocytes (Fig. 4d). These data suggest that T39 facilitates LXRα polyubiquitination and turnover, and that, in *T39*^{-/-} cells, LXRα is protected from proteasomal degradation. We found a significant inverse correlation between published data on LXRα binding strength to active regulatory elements in the liver²⁵ and the magnitude of target gene induction in T39 deficiency relative to WT (Extended Data Fig. 8d), suggesting that the impact of increased LXR protein levels is larger for genes containing lower-affinity LXR binding sites, such as *Insig2a* and *Lpcat3*, and minimally affects genes containing high-affinity LXR binding sites, including *Srebf1* and *Fasn*. Hepatic LXR overexpression has been reported to upregulate cholesterol removal genes such as *Abcg5/8* (ref. 26), which is consistent with our findings. However, there were minimal effects on SREBP-1 target genes²⁶, suggesting that decreased LXR ubiquitination may also contribute to the decreased lipogenesis in *T39*^{-/-} mice.

The deficiency of T39 demonstrates that, as opposed to the effects of potent synthetic ligands, decreasing the ubiquitination and increasing the abundance of endogenous LXR protein can activate anti-atherogenic cholesterol removal while inhibiting the lipogenesis that leads to steatosis (Extended Data Fig. 9). By ameliorating cardiovascular disease and non-alcoholic fatty liver disease, T39 inhibition could offer a new therapeutic approach to tackle two globally prevalent chronic diseases.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 December 2013; accepted 2 June 2016.

Published online 6 July 2016.

- Musso, G., Gambino, R., Cassader, M. & Pagano, G. A meta-analysis of randomized trials for the treatment of nonalcoholic fatty liver disease. *Hepatology* **52**, 79–104 (2010).
- Teslovich, T. M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713 (2010).
- Willer, C. J. *et al.* Discovery and refinement of loci associated with lipid levels. *Nature Genet.* **45**, 1274–1283 (2013).
- McNeish, J. *et al.* High density lipoprotein deficiency and foam cell accumulation in mice with targeted disruption of ATP-binding cassette transporter-1. *Proc. Natl Acad. Sci. USA* **97**, 4245–4250 (2000).
- Brunham, L. R. *et al.* Intestinal ABCA1 directly contributes to HDL biogenesis *in vivo*. *J. Clin. Invest.* **116**, 1052–1062 (2006).
- Wu, A. L. & Windmueller, H. G. Relative contributions by liver and intestine to individual plasma apolipoproteins in the rat. *J. Biol. Chem.* **254**, 7316–7322 (1979).
- Costet, P., Luo, Y., Wang, N. & Tall, A. R. Sterol-dependent transactivation of the ABC1 promoter by the liver X receptor/retinoid X receptor. *J. Biol. Chem.* **275**, 28240–28245 (2000).
- Zelcer, N., Hong, C., Boyadjian, R. & Tontonoz, P. LXR regulates cholesterol uptake through Idol-dependent ubiquitination of the LDL receptor. *Science* **325**, 100–104 (2009).
- Matsuzawa, N. *et al.* Lipid-induced oxidative stress causes steatohepatitis in mice fed an atherogenic diet. *Hepatology* **46**, 1392–1403 (2007).
- Rong, X. *et al.* LXRs regulate ER stress and inflammation through dynamic modulation of membrane phospholipid composition. *Cell Metab.* **18**, 685–697 (2013).
- Bensinger, S. J. *et al.* LXR signaling couples sterol metabolism to proliferation in the acquired immune response. *Cell* **134**, 97–111 (2008).
- Lo Sasso, G. *et al.* Intestinal specific LXR activation stimulates reverse cholesterol transport and protects from atherosclerosis. *Cell Metab.* **12**, 187–193 (2010).
- Zhang, Y. *et al.* Liver LXR α expression is crucial for whole body cholesterol homeostasis and reverse cholesterol transport in mice. *J. Clin. Invest.* **122**, 1688–1699 (2012).
- Yu, L. *et al.* Overexpression of ABCG5 and ABCG8 promotes biliary cholesterol secretion and reduces fractional absorption of dietary cholesterol. *J. Clin. Invest.* **110**, 671–680 (2002).
- Repa, J. J. *et al.* Regulation of mouse sterol regulatory element-binding protein-1c gene (SREBP-1c) by oxysterol receptors, LXR α and LXR β . *Genes Dev.* **14**, 2819–2830 (2000).
- Hegarty, B. D. *et al.* Distinct roles of insulin and liver X receptor in the induction and cleavage of sterol regulatory element-binding protein-1c. *Proc. Natl Acad. Sci. USA* **102**, 791–796 (2005).
- Huang, Y. *et al.* A feed-forward loop amplifies nutritional regulation of PNPLA3. *Proc. Natl Acad. Sci. USA* **107**, 7892–7897 (2010).
- Romeo, S. *et al.* Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nature Genet.* **40**, 1461–1465 (2008).
- Vatner, D. F. *et al.* Insulin-independent regulation of hepatic triglyceride synthesis by fatty acids. *Proc. Natl Acad. Sci. USA* **112**, 1143–1148 (2015).
- Walker, A. K. *et al.* A conserved SREBP-1/phosphatidylcholine feedback circuit regulates lipogenesis in metazoans. *Cell* **147**, 840–852 (2011).
- Hishikawa, D. *et al.* Discovery of a lysophospholipid acyltransferase family essential for membrane asymmetry and diversity. *Proc. Natl Acad. Sci. USA* **105**, 2830–2835 (2008).
- Rong, X. *et al.* Lpcat3-dependent production of arachidonoyl phospholipids is a key determinant of triglyceride secretion. *eLife* **4**, <http://dx.doi.org/10.7554/eLife.06557> (2015).
- Hannah, V. C., Ou, J., Luong, A., Goldstein, J. L. & Brown, M. S. Unsaturated fatty acids down-regulate srebp isoforms 1a and 1c by two mechanisms in HEK-293 cells. *J. Biol. Chem.* **276**, 4365–4372 (2001).
- Kammoun, H. L. *et al.* GRP78 expression inhibits insulin and ER stress-induced SREBP-1c activation and reduces hepatic steatosis in mice. *J. Clin. Invest.* **119**, 1201–1215 (2009).
- Boergesen, M. *et al.* Genome-wide profiling of liver X receptor, retinoid X receptor, and peroxisome proliferator-activated receptor α in mouse liver reveals extensive sharing of binding sites. *Mol. Cell. Biol.* **32**, 852–867 (2012).
- Lehrke, M. *et al.* Diet-dependent cardiovascular lipid metabolism controlled by hepatic LXR α . *Cell Metab.* **1**, 297–308 (2005).

Supplementary Information is available in the online version of the paper.

Acknowledgements We express our gratitude to F. Matsuura for support, A. Morishita for advice on liver histology, M. Sakurai and T. Yamashita for advice on immunoprecipitation experiments, W. R. Lagor for advice on the reverse cholesterol transport study, M. Ishibashi for advice on animal administration, J. W. Medley for consultation on the coupling reaction, and N. Wang for project discussions. D. J. Gorman and J. So provided technical support, and O. Xu provided technical services for the lipidomics analysis. This work was supported by grants from the Manpei Suzuki Diabetes Foundation (to M.K.), VIDI grant 91715350 from the Netherlands Organization of Sciences (to M.W.), Rosalind Franklin Fellowship from the University Medical Center Groningen (to M.W.), JSPS KAKENHI Grant 15K160203 (to I.I.), and the Fondation Leducq (to A.R.T.). This work was supported by grants from the National Institutes of Health (T32 training program HL007343, M.M.M.; HL087123 and HL119830, to A.R.T.; HL101864 and HL111398, to D.J.R.; DK46900, to M.M.H.).

Author Contributions J.H. and M.K. generated epitope-tagged constructs, bred mice, performed *in vivo* and cell culture experiments, collected data, designed the study, interpreted data, and wrote the paper; M.M.M. isolated hepatocytes and performed the chromatin immunoprecipitation and insulin sensitivity experiments; E.Y. and L.T. collected data; I.I. performed oxysterol and plant sterol measurements; M.W., S.A. and C.B.W. performed atherosclerotic lesion analysis; R.B.C. and G.D. designed and performed lipidomics analyses; J.I. performed enterocyte *ex vivo* secretion studies; S.T. performed microarray analysis; J.H.L. performed histopathological analyses of liver sections; D.J.R., M.M.H., and S.Y. were involved in study design; A.R.T. designed the study, interpreted data, and wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.R.T. (art1@cumc.columbia.edu) or M.K. (koseki@cardiology.med.osaka-u.ac.jp).

Reviewer Information *Nature* thanks C. Semenkovich, A. von Eckardstein and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Mice and diet. Both male and female 10- to 20-week-old littermate mice were used in this study, and mice were allocated to experiments to match age and sex. $T39^{-/-}$ mice on the C57Bl/6N background were obtained from the Wellcome Trust Sanger Institute (Hinxton, UK). $Ldlr^{-/-}$ $T39^{-/-}$ mice were generated by crossing these mice with $Ldlr^{-/-}$ mice from Jackson Laboratory (Bar Harbour, Maine). $LXR\alpha^{-/-}$ $T39^{-/-}$ mice were generated by crossing $T39^{-/-}$ mice with $LXR\alpha^{-/-}$ mice from Jackson Laboratory. For all backgrounds, $T39^{+/+}$ and $T39^{-/-}$ mice were generated by breeding $T39^{+/+}$ mice. Tissue-specific $T39$ knockout mice were generated by crossing $T39^{fl/fl}$ mice from Merck/TaconicArtemis (Germany) with *Villin-Cre* and *Albumin-Cre* mice obtained from the Jackson Laboratory. At 8 weeks of age, mice were started on either non-irradiated HF/HC/BS diet (7.5% cocoa butter, 1.25% cholesterol, 0.5% sodium cholate, no. 88051) or irradiated WTD (21% milk fat, 0.15% cholesterol, no. 88137) (Harlan Teklad). Mice were housed in a specific pathogen-free facility on a 12h:12h light:dark cycle. Compatible mice of mixed genotypes were housed in groups of five. Each mouse was assigned a unique identification number that did not indicate genotype, and experiments and measurements were conducted in a blinded manner until data analysis. Animal numbers were selected on the basis of power calculations of 0.8 using variances from previous studies and availability of genotypes that arose from heterozygote breedings. Protocols were approved by the Institutional Animal Care and Use Committee of Columbia University.

Lipoprotein analysis. For chow and HF/HC/BS diet-fed samples, HDL cholesterol was determined in the supernatant of phosphotungstate/ Mg^{2+} -precipitated serum using an enzymatic-based colorimetric assay (Wako). Enzymatic-based colorimetric assays were also used to determine total serum cholesterol (Wako) and triglyceride (Infinity Triglycerides, Thermo Scientific). For WTD-fed $Ldlr^{-/-}$ mice, lipoproteins were separated by KBr density ultracentrifugation and then assayed using the above kits.

Enterocyte collection, protein extraction, and western blotting. Enterocytes were collected from the jejunum using Cell Recovery Solution (BD) in a method adapted from Perreault & Beaulieu²⁷. Enterocytes and liver samples were homogenized in 1% Triton X-100, 0.5% sodium deoxycholate, 2 mM EDTA, and 1 mM EGTA in phosphate buffered saline, pH 7.4. Cytoplasmic and nuclear fractions were extracted from hepatocytes using the CellLytic NuCLEAR Extraction kit (Sigma-Aldrich) according to the manufacturer's instructions. Proteins were resolved by SDS-PAGE and transferred to a polyvinylidene fluoride membrane for western blotting. Western blotting antibodies were rabbit anti- $LXR\alpha/\beta$ (H-144, Santa Cruz Biotechnology), mouse anti- $LXR\alpha$ (PPZ0412, Abcam), rabbit anti-ABCA1 (NB400-105, Novus Biologicals), mouse anti-SREBP-1 (clone 2A4, Thermo Scientific), rabbit anti-HSP90 (C45G5, Cell Signaling), goat anti-histone H3 (ab12079, Abcam), mouse anti-polyubiquitinated proteins (clone FK2, Millipore), and rabbit anti-SREBP-2 (14508-1-AP, Proteintech Group). An affinity-purified antibody against mouse and human $T39$ was generated by immunizing rabbits for 118 days against a conjugated amino-acid sequence of $T39$, NH_2 -CKESKWSKATYVFLKAAILS-COOH (Covance). The integrated density value of immunoblot signals was quantified by AlphaEaseFC software (Alpha Innotech).

Secretion of [3H]cholesterol by primary enterocytes. For characterization of secreted lipoproteins, enterocytes were isolated from overnight fasted mice, and radiolabelled for 1 h with $0.5 \mu Ci/ml$ of [3H]cholesterol, washed, and incubated with fresh media containing lipid/bile salt micelles consisting of 1.4 mM oleic acid, 0.14 mM sodium cholate, 0.15 mM sodium deoxycholate, 0.17 mM phosphatidylcholine, and 0.19 mM mono-oleoylglycerol^{28,29}. After 2 h, enterocytes were centrifuged and supernatants were collected. Media were subjected to density gradient ultracentrifugation to determine radiolabelled cholesterol distribution among lipoprotein classes²⁹.

Grading of non-alcoholic steatohepatitis. Frozen sections of liver were stained with haematoxylin and eosin. Livers from mice fed HF/HC/BS diet were assessed by a blinded observer at $200\times$ magnification according to the staging outlined by Kleiner *et al.*³⁰. The lobular inflammation grade was assigned a score from 0 to 3 as follows: no foci = 0; fewer than two foci per field = 1; two to four foci per field = 2; more than four foci per field = 3. Liver cell injury was assigned a score from 0 to 2 as follows: no ballooned cells = 0; a few ballooned cells = 1; many ballooned cells = 2; majority of hepatocytes are ballooned = 3. The fibrosis grade was assigned a score from 0 to 3 as follows: no fibrosis = 0; scattered centrilobular and perisinusoidal areas of fibrosis = 1; many regions of fibrosis and true bridging = 2; extensive fibrosis and bridging, with/without cirrhosis = 3.

Dietary cholesterol absorption. Mice were fasted for 14 h and then given an intraperitoneal injection of Poloxamer-407 (1,000 mg/kg, Sigma-Aldrich) to inhibit peripheral triglyceride-rich lipoprotein catabolism. The mice were administered $10 \mu Ci$ [3H]cholesterol (PerkinElmer) in $135 \mu L$ olive oil containing 0.1 mg cold cholesterol, and blood was sampled from the tail vein every 4 h. To inhibit NPC1L1,

mice were gavaged with ezetimibe (10 mg/kg, Selleckchem) once daily for 3 days before the experiment.

Sterol analysis and GC-MS analysis. Oxysterols were measured as previously described³¹. 19-Hydroxycholesterol (Steraloids) was added to samples as an internal standard. Lipids were extracted from plasma ($50 \mu L$) or livers (50 mg wet weight) with chloroform/methanol/water (1:2:0.8, v/v) containing butylated hydroxytoluene according to the Bligh and Dyer method³². Then the lipids were saponified at room temperature overnight in the dark. Unsaponified lipid was applied to a Sep-Pak Vac silica cartridge (Waters) to separate oxysterols and sterols³³. Trimethylsilyl derivatives of the sterols were quantified by gas chromatography-mass spectrometry using a GC/MS QP2010 (Shimadzu) equipped with a SPB-1 fused silica capillary column ($60 m \times 0.25 mm \times 0.25 \mu m$; Spelco). In the oven temperature programme, the temperature was initiated at $180^\circ C$ for 1 min and then raised to $250^\circ C$ at $20^\circ C/min$ and to $290^\circ C$ at $5^\circ C/min$ then held for 45 min. The injection temperature was set at $300^\circ C$, the interface at $300^\circ C$, and the ion source adjusted to $200^\circ C$.

In vivo determination of de novo lipogenesis. WTD-fed mice were fasted overnight and then refed for 4 h. Mice were given an intraperitoneal injection of $1.5 mCi$ [3H]water, and the liver was harvested 2 h later. Hepatic lipids were extracted by the Folch method and resolved by thin layer chromatography. The lipid spots visualized by iodine were scraped and 3H incorporation was determined by liquid scintillation counting. Biosynthetic rate was calculated essentially as described previously by Spady and Dietsch³⁴.

Lipidomics analysis. Microsomes were isolated from livers from HF/HC/BS diet-fed mice. Briefly, livers were homogenized with a Dounce homogenizer in a KCl-sucrose buffer (0.05 M KH_2PO_4 , 0.25 M sucrose, 0.154 M KCl, pH 7.5) supplemented with protease inhibitors. Cellular debris, mitochondria, nuclei and plasma membrane were removed with a 10,000g centrifugation, and microsomes were pelleted with a 100,000g ultracentrifugation. Lipids were extracted by the Folch method in chloroform:methanol supplemented with butylated hydroxytoluene. Lipid extracts of purified microsome were spiked with a cocktail of internal standards, and analysed using a 6490 Triple Quadrupole LC/MS system (Agilent Technologies). Glycerophospholipids and sphingolipids were separated with normal-phase HPLC as described before³⁵, with a few modifications. A Zorbax Rx-Sil column (inner diameter $2.1 mm \times 100 mm$, Agilent) was used under the following conditions: mobile phase A (chloroform:methanol:1 M ammonium hydroxide, 89.9:10.0:1, v/v) and mobile phase B (chloroform:methanol:water: ammonium hydroxide, 55:39.9:5.0:1, v/v); 95% A for 2 min, linear gradient to 30% A over 18 min and held for 3 min, and linear gradient to 95% A over 2 min and held for 6 min. Quantification of lipid species was accomplished using multiple reaction monitoring transitions and instrument settings that were determined in earlier studies³⁵ in conjunction with referencing of known amounts of internal standards: phosphatidic acid 14:0/14:0, phosphatidylcholine 14:0/14:0, phosphatidylethanolamine 14:0/14:0, phosphatidylinositol 12:0/13:0, phosphatidylserine 14:0/14:0, sphingomyelin d18:1/12:0 (Avanti Polar Lipids).

Atherosclerosis study. $Ldlr^{-/-}$ and $Ldlr^{-/-}$, $T39^{-/-}$ mice were fed WTD for 20 weeks. Mice were killed in accordance to the American Veterinary Association Panel. Hearts and aortas were perfused with phosphate buffered saline, isolated and fixed in neutral phosphate-buffered formalin. The aortic arch and the descending aorta were stained with Oil Red O. Aortas were pinned on silicon dishes and Oil Red O positive areas were quantified using Image J software and expressed as the percentage of the total aorta area. Hearts were dehydrated, embedded in paraffin, and the aortic root area was cross-sectioned in $5 \mu m$ sections. The sections were stained with haematoxylin and eosin and the average of six evenly distributed sections for each animal was used to determine lesion size and complexity. Lesion size was quantified by morphometric analysis using Image-Pro Plus software (Media Cybernetics). The typing of lesions is done according to the typing for humans proposed by the American Heart Association³⁶ and adapted to categorize murine lesions³⁷. In this study, we discerned sections showing macrophage foam cell rich lesions (type I-II), complex lesions with fibrous caps (type III), and advanced lesions with foam cells in the media and presence of fibrosis, cholesterol clefts, mineralization and/or necrosis (type IV-V).

Plasmids. Human $T39$ was cloned from HepG2 cDNA into pCMV6-AN-GFP, pCMV6-Entry, and pCMV6-AN-Myc-DDK vectors (Origene) linearized by SgfI and MluI digestion using the In-Fusion PCR cloning kit (Clontech). Using the same approach, murine $Nr1h3$ cDNA in the pCMV-SPORT6 vector (DF/HCC DNA Resource Core, Harvard University) was subcloned into the pCMV6-Entry and pCMV6-AN-Myc-DDK vectors.

Primary hepatocyte culture and transfection. Hepatocytes were isolated from 6- to 12-week-old littermate mice using the two-step perfusion method and seeded onto collagen I-coated dishes (BD Bionocoat) at a density of 2.5×10^4 cells/cm² in DMEM (Cellgro) containing 10% fetal bovine serum (Gibco) and 1% penicillin/streptomycin (Gibco). Cells were crosslinked and harvested for chromatin

immunoprecipitation analysis 2 h following hepatocyte isolation. Hepatocytes were transfected the following day using JetPEI-Gal (Polyplus) according to the manufacturer's instructions and experiments were performed 3 days after transfection.

Cell culture and transfection. HEK293T cells (ATCC) were propagated in DMEM containing 10% fetal bovine serum and 1% penicillin/streptomycin. Cells were transfected with Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Cells were not tested for mycoplasma contamination after receipt from the ATCC.

Chromatin immunoprecipitation. Mouse hepatocytes were harvested and cultured as described above. At the time of plating, hepatocytes were treated with either DMSO or 2 μ M GW3965. Two hours after treatment, hepatocytes were crosslinked in 1% formaldehyde, at room temperature, for 10 or 20 min (depending on IP). The cross-linking reaction was quenched using 0.125 M glycine (10 min at room temperature). Cells were washed twice with cold PBS, then scraped in cold PBS and pelleted at 3,000 r.p.m. The pellet was then resuspended in 1% SDS sonication buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA) and sonicated using a Fisher Scientific 550 Sonic Dismembrator. Using a cycle setting of 7, lysates were sonicated for 15 s followed by a 30 s rest; two sets of six pulses were performed to acquire sheared DNA between 200 and 1,000 bp. Lysates were then cleared at 14,000 r.p.m. for 10 min at 4°C; 500 μ g of chromatin was used for each IP. Antibodies used for chromatin immunoprecipitation: anti-RXR (Δ N 197, Santa Cruz Biotechnology) and Rabbit IgG (sc2027, Santa Cruz Biotechnology). Primers for chromatin immunoprecipitation analysis were designed according to LXREs identified by ChipSeq²⁵.

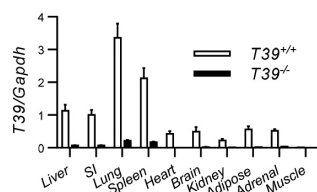
Isolation of endogenous LXR with GW3965-affinity beads. Female mice fed HF/HC/BS diet were fed HF/HC/BS diet for 6 weeks, and infected with CMV-Ubc adenovirus (Vector Biolabs) during the last week. Mice were given an oral gavage of the proteasomal inhibitor ixazomib (20 mg/kg in 40% 2-hydroxypropyl- β -cyclodextrin, Selleckchem) and then fasted for 6 h before being killed. GW3965 HCl (Selleckchem) was coupled to M-270 amine Dynabeads (Invitrogen) in a nitrogen-purged reaction containing *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide, 4-(dimethylamino)pyridine, and tributylamine in dimethylformamide (Sigma). Replacing the carboxylic acid in GW3965 with an amide in the coupling reaction allows the synthetic LXR agonist to retain its ligand binding domain recruitment activity³⁸. Livers from HF/HC/BS diet-fed mice were sonicated in HEPES lysis buffer (50 mM HEPES, 150 mM NaCl, 1% NP-40, 10% glycerol, 5 mM EDTA, pH 7.4) supplemented with 10 mM *N*-ethylmaleimide, 10 mM iodoacetamide (Sigma), 25 μ M PR-619 (Calbiochem), 5 μ M MG-132 (Selleckchem), TUBE-2 (LifeSensors), and Halt Protease Inhibitor Cocktail (Thermo Scientific). Liver lysates were incubated with GW3965-beads at 37°C for 2 h.

LXR α protein turnover. Flag-Myc-LXR α -transfected primary hepatocytes were cultured in methionine-free media for 1 h, and then protein synthesis was labelled with 400 μ M L-homopropargylglycine (HPG) for 30 min. The labelled primary hepatocytes were washed and then cultured in 20 mM L-methionine-enriched

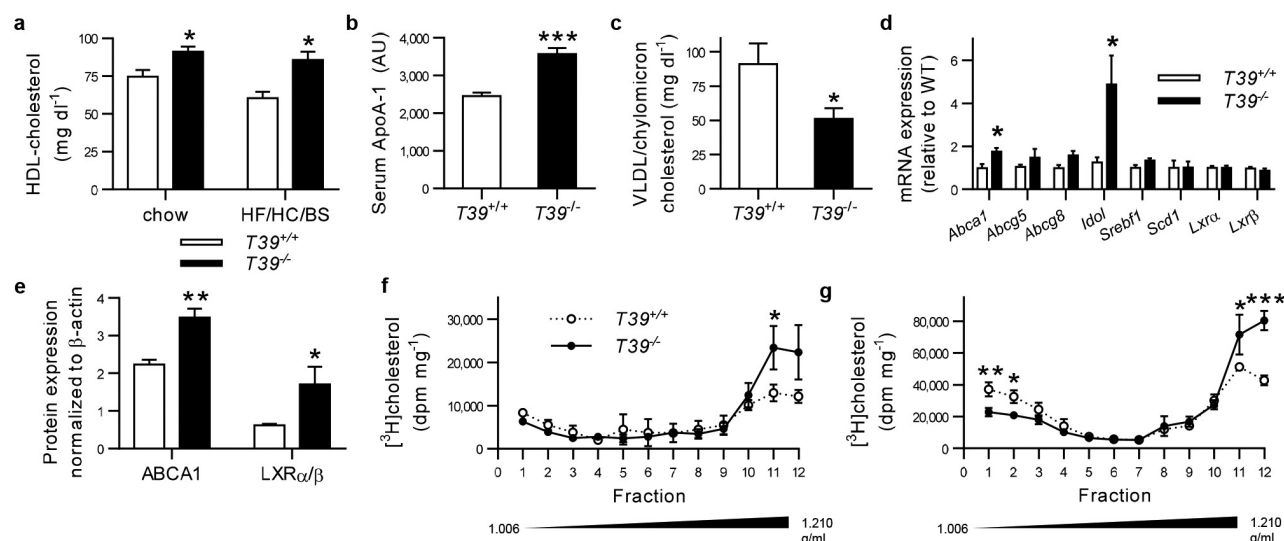
media with/without bortezomib for the duration of the chase. Flag-Myc-LXR α was immunoprecipitated with anti-Flag (clone M2)-conjugated magnetic beads (Sigma-Aldrich) and HPG-labelled proteins were conjugated to tetramethylrhodamine azide by click chemistry (Molecular Probes) before separation by SDS-PAGE for visualization.

Statistical analysis. No statistical methods were used to predetermine sample size. The experiments were not randomized. All data are represented as mean \pm s.e.m. and, unless otherwise stated, were analysed using a two-tailed Student's *t*-test or two-way ANOVA with Bonferroni post hoc analysis where appropriate. In experiments where $n \geq 8$ for all genotypes, the D'Agostino-Pearson omnibus test for normality and Bartlett's test for variance were used to ensure that the assumptions of the statistical tests were met. All tests were performed with Prism (Graphpad) and $P < 0.05$ was considered statistically significant. Values/animals were excluded if it was detected as a significant ($P < 0.05$) outlier on the basis of a two-sided Grubbs' test.

27. Perreault, N. & Beaulieu, J. F. Primary cultures of fully differentiated and pure human intestinal epithelial cells. *Exp. Cell Res.* **245**, 34–42 (1998).
28. Anwar, K., Iqbal, J. & Hussain, M. M. Mechanisms involved in vitamin E transport by primary enterocytes and in vivo absorption. *J. Lipid Res.* **48**, 2028–2038 (2007).
29. Iqbal, J., Anwar, K. & Hussain, M. M. Multiple, independently regulated pathways of cholesterol transport across the intestinal epithelial cells. *J. Biol. Chem.* **278**, 31610–31620 (2003).
30. Kleiner, D. E. *et al.* Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* **41**, 1313–1321 (2005).
31. Tomoyori, H. *et al.* Phytosterol oxidation products are absorbed in the intestinal lymphatics in rats but do not accelerate atherosclerosis in apolipoprotein E-deficient mice. *J. Nutr.* **134**, 1690–1696 (2004).
32. Bligh, E. G. & Dyer, W. J. A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* **37**, 911–917 (1959).
33. Dzeletovic, S., Breuer, O., Lund, E. & Diczfalussy, U. Determination of cholesterol oxidation products in human plasma by isotope dilution-mass spectrometry. *Anal. Biochem.* **225**, 73–80 (1995).
34. Spady, D. K. & Dietschy, J. M. Sterol synthesis in vivo in 18 tissues of the squirrel monkey, guinea pig, rabbit, hamster, and rat. *J. Lipid Res.* **24**, 303–315 (1983).
35. Chan, R. B. *et al.* Comparative lipidomic analysis of mouse and human brain with Alzheimer disease. *J. Biol. Chem.* **287**, 2678–2688 (2012).
36. Stary, H. C. *et al.* A definition of advanced types of atherosclerotic lesions and a histological classification of atherosclerosis. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association. *Circulation* **92**, 1355–1374 (1995).
37. Gijbels, M. J. *et al.* Progression and regression of atherosclerosis in APOE3-Leiden transgenic mice: an immunohistochemical study. *Atherosclerosis* **143**, 15–25 (1999).
38. Collins, J. L. *et al.* Identification of a nonsteroidal liver X receptor agonist through parallel array synthesis of tertiary amines. *J. Med. Chem.* **45**, 1963–1966 (2002).

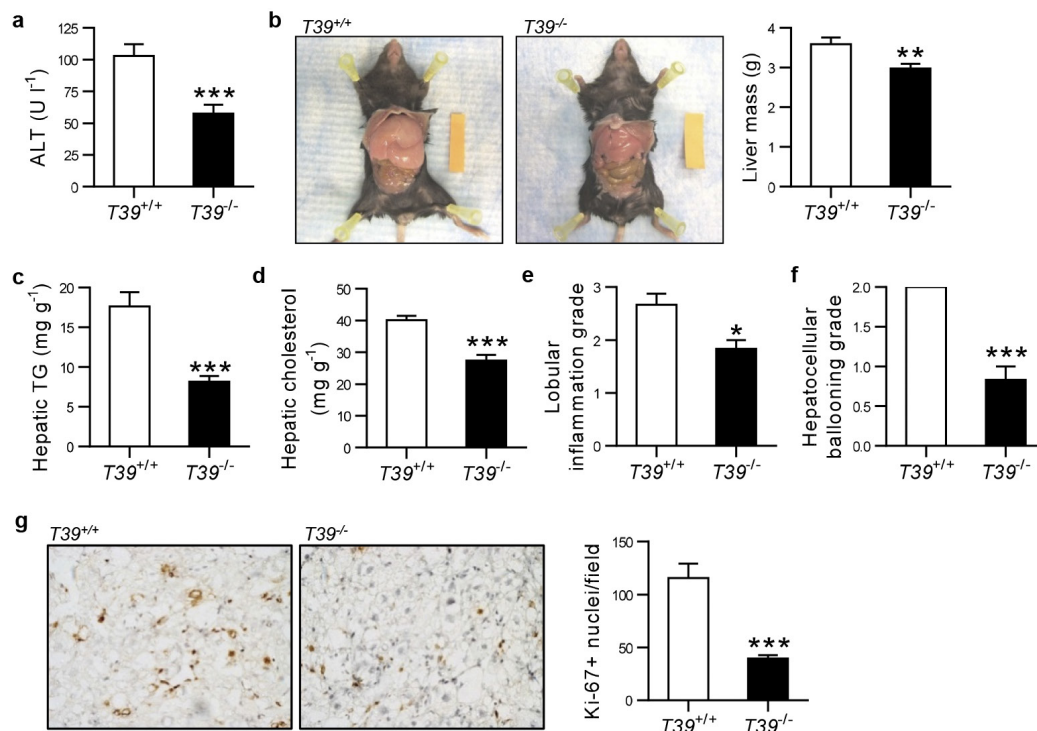
**Extended Data Figure 1 | Organ distribution of T39 mRNA expression.**

T39 mRNA levels were assessed in chow-fed male WT and whole-body *T39* knockout mice. SI, small intestine; $n = 4$ animals per genotype.



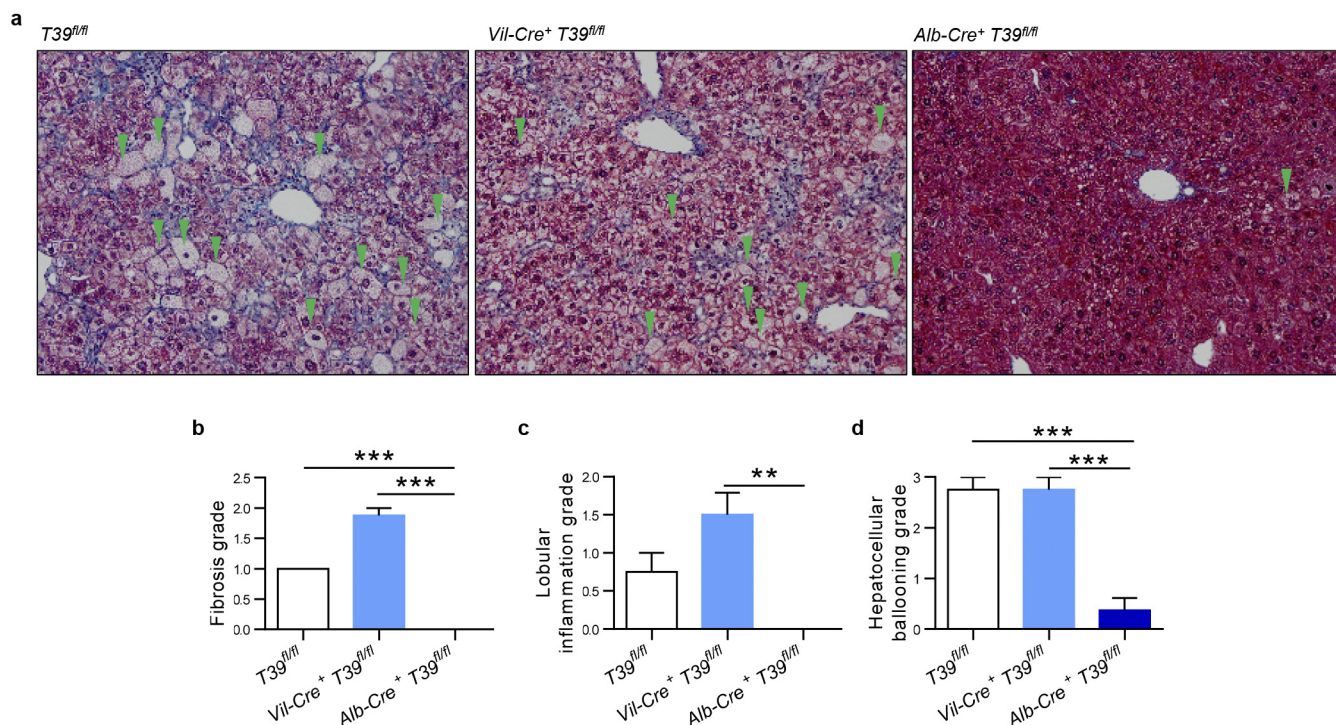
Extended Data Figure 2 | Beneficial lipoprotein changes in T39-deficient mice fed HF/HC/BS diet. **a**, HDL cholesterol of WT and whole-body T39 knockout mice fed chow or HF/HC/BS diet for 4 weeks, $n = 4$ per genotype for chow, 5 per genotype for HF/HC/BS diet. **b**, Serum ApoA-1 as determined by SDS-PAGE and **(c)** serum VLDL/chylomicron cholesterol of WT and whole-body T39 knockout mice fed HF/HC/BS diet for 4 weeks, $n = 5$ per genotype. **d**, Enterocyte mRNA expression, $n = 5$ per genotype. **e**, Quantification of enterocyte protein expression normalized to β -actin, $n = 4$ per genotype. **f**, Cholesterol secretion profiles of enterocytes

collected from WT and whole-body T39 knockout mice fed chow or **(g)** HF/HC/BS diet. Enterocytes were incubated with taurocholate micelles containing [³H]cholesterol for 2 h, and then secreted lipoproteins were separated by density ultracentrifugation. Lipoprotein fractions are shown with fraction 1 being the most buoyant and fraction 12 the most dense, $n = 3$ per genotype, and replicated in two different experiments. For all panels, data are represented as mean \pm s.e.m., * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$ by two-tailed Student's t -test.



Extended Data Figure 3 | Improved features of non-alcoholic steatohepatitis in *T39*^{-/-} mice fed HF/HC/BS diet. WT and whole-body *T39* knockout mice were fed the HF/HC/BS diet for 18 weeks. **a**, Serum ALT, **(b)** liver size, **(c)** hepatic triglyceride, and **(d)** hepatic cholesterol content, $n = 15$ – 19 per genotype. Grading of **(e)** inflammatory cell infiltration and **(f)** hepatocellular ballooning degeneration on the basis

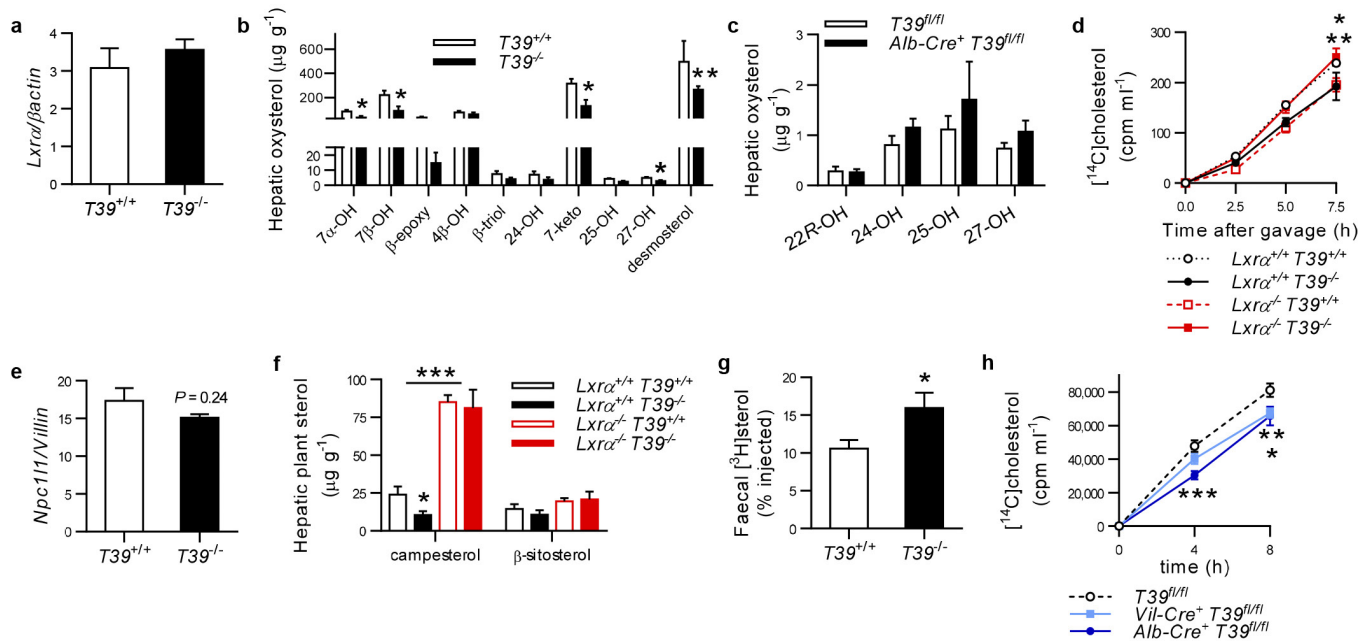
of haematoxylin and eosin-stained sections, $n = 6$ per genotype. **g**, Ki-67 was immunohistochemically detected with diaminobenzidine (brown) in frozen liver sections with nuclei counterstained with haematoxylin (blue). Quantification of Ki-67-positive nuclei is shown on the right, mean of five fields, $n = 6$ animals/genotype. Data are represented as mean \pm s.e.m., ** $P < 0.01$ and *** $P < 0.001$ by two-tailed t -test.



Extended Data Figure 4 | Improvements in the histological features of non-alcoholic steatohepatitis in liver-specific T39 knockout.

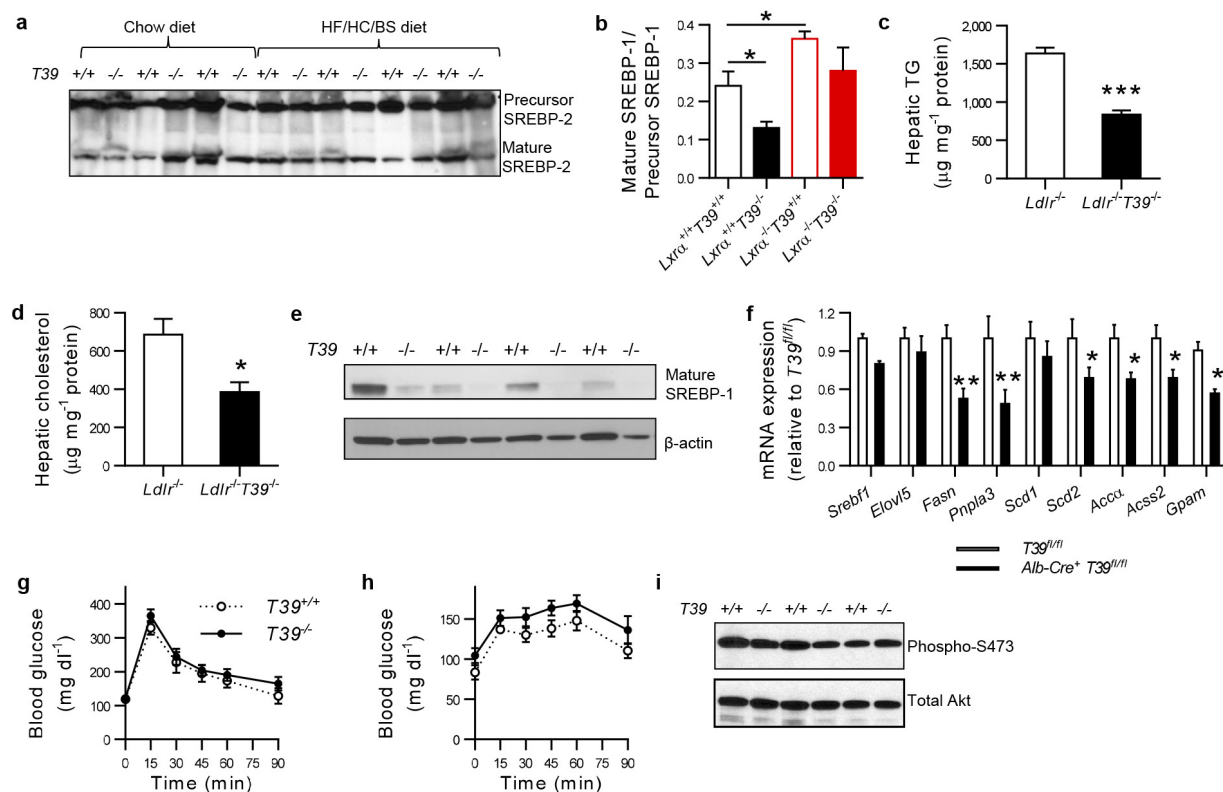
Enterocyte- and hepatocyte-specific T39 knockout and $T39^{fl/fl}$ control mice were fed HF/HC/BS diet for 21 weeks. **a**, Representative Masson's trichrome stain of hepatic sections at 100 \times magnification, showing cytoplasm as red, collagen as blue, and nuclei as dark brown. Green

arrowheads point to ballooned hepatocytes. Grading of **(b)** fibrosis severity, **(c)** inflammatory foci number and **(d)** extent of hepatocellular ballooning on the basis of Masson's trichrome staining, $n = 4$ per genotype. Data are represented as mean \pm s.e.m., ** $P < 0.01$, *** $P < 0.001$ by one-way ANOVA.



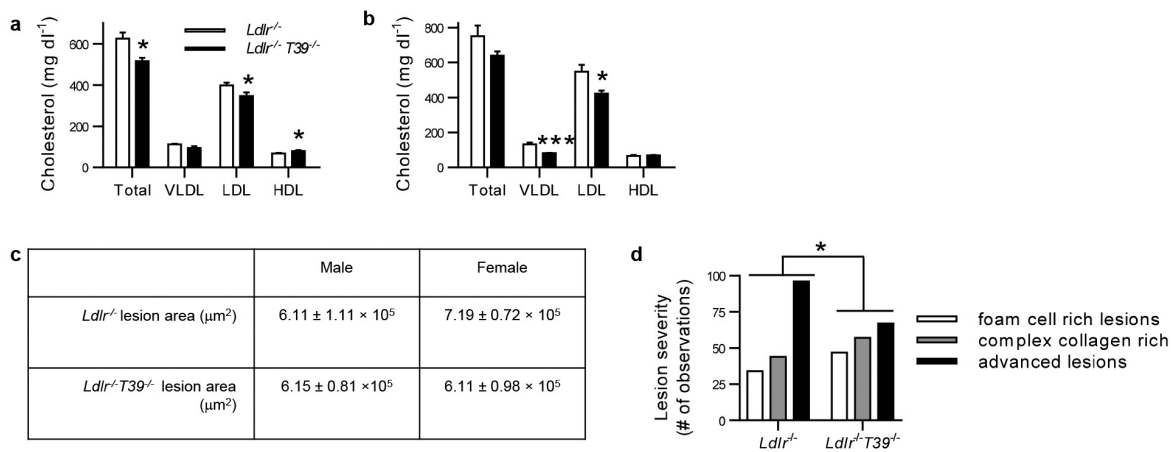
Extended Data Figure 5 | Decreased dietary cholesterol absorption in T39-deficient mice fed HF/HC/BS diet. **a**, Hepatic *Lxrα* mRNA expression in mice fed HF/HC/BS diet for 5 weeks, *n* = 4 per genotype. **b**, Hepatic oxysterol content of WT and whole-body T39 knockout mice fed HF/HC/BS diet for 18 weeks, *n* = 5 WT and 4 *T39*^{-/-}. **c**, Hepatic content of endogenous LXR ligands in control and liver-specific T39 knockout mice fed HF/HC/BS diet, *n* = 7 per genotype. **d**, Absorption of [¹⁴C]cholesterol administered by gavage to WT and whole-body T39 knockout with/without LXRα along with Poloxamer-407 injection to inhibit peripheral lipoprotein catabolism, *n* = 3 *Lxrα*^{+/+} *T39*^{+/+}, 4 *Lxrα*^{+/+} *T39*^{-/-}, 3 *Lxrα*^{-/-} *T39*^{+/+}, and 5 *Lxrα*^{-/-} *T39*^{+/+}. **e**, Enterocyte *Npc1l1* mRNA expression in WT and whole-body T39 knockout mice,

n = 5 per genotype. **f**, Hepatic plant sterol content of WT and whole-body T39 knockout mice with/without LXRα fed HF/HC/BS diet for 6 weeks, *n* = 4 *Lxrα*^{+/+} *T39*^{+/+}, 4 *Lxrα*^{+/+} *T39*^{-/-}, 4 *Lxrα*^{-/-} *T39*^{+/+}, and 6 *Lxrα*^{-/-} *T39*^{+/+}. **g**, Reverse cholesterol transport of WT and whole-body T39 knockout mice fed HF/HC/BS diet. [³H] faecal excretion was measured over 3 days after an intravenous injection of [³H]cholesteryl ester-labelled HDL, *n* = 7 WT and 5 *T39*^{-/-}. **h**, Absorption of [¹⁴C]cholesterol administered by gavage to tissue-specific T39 knockout mice fed HF/HC/BS diet for 5 weeks and injected with Poloxamer-407, *n* = 8 *T39*^{fl/fl}, 7 *Vill-Cre*⁺ *T39*^{fl/fl}, and 9 *Albumin-Cre*⁺ *T39*^{fl/fl}. Data are represented as mean ± s.e.m., **P* < 0.05, ***P* < 0.01, ****P* < 0.001 by two-tailed *t*-test or two-way ANOVA for absorption studies.



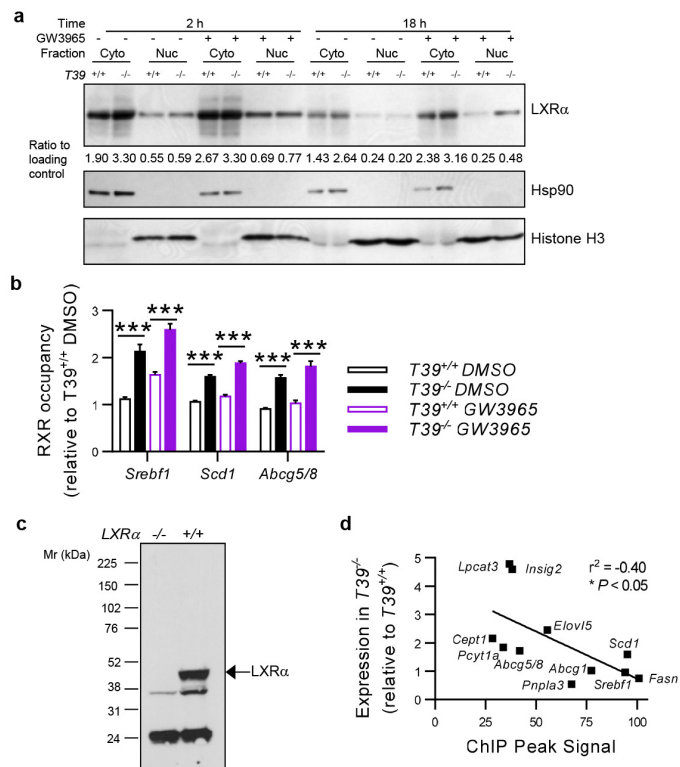
Extended Data Figure 6 | Decreased SREBP-1 processing without altered hepatic insulin sensitivity in T39-deficient mice. **a**, Hepatic protein expression of precursor SREBP-2 (upper band) and processed SREBP-2 (lower band) of WT and whole-body T39 knockout mice fed HF/HC/BS diet for 5 weeks. **b**, SREBP-1 processing ratio in WT and whole-body T39 knockout mice with/without LXR α fed HF/HC/BS diet for 6 weeks as based on quantification of an anti-SREBP-1 immunoblot, $n = 4$ per genotype. **c**, Hepatic triglyceride and **(d)** cholesterol content of WT and whole-body T39 knockout mice lacking LDLR fed WTD for 20 weeks, $n = 5$ per genotype. **e**, Protein expression of the nuclear form of SREBP-1 (top) and β -actin loading control (bottom) of WTD-fed WT and whole-body T39 knockout mice lacking LDLR fed WTD for 20 weeks. **f**, Hepatic

gene expression in liver-specific T39 knockout and control animals fed HF/HC/BS diet for 18 weeks, following a fasting/refeeding protocol, $n = 13$ T39^{fl/fl} and 7 Albumin-Cre⁺ T39^{fl/fl}. **g**, Intraperitoneal glucose tolerance of WT and whole-body T39 knockout mice fed HF/HC/BS diet after a 6 h fast, $n = 8$ WT and 12 T39^{-/-}. **h**, Pyruvate tolerance of WT and whole-body T39 knockout mice fed HF/HC/BS diet after an overnight fast, $n = 5$ WT and 7 T39^{-/-}. **i**, Hepatic Akt phosphorylation 5 min after portal vein delivery of insulin in WT or whole-body T39 knockout mice fed HF/HC/BS diet. Data are represented as mean \pm s.e.m., * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ by two-tailed Student's t -test. See Supplementary Fig. 1 for gel source data.



Extended Data Figure 7 | Improved lipoprotein profile and less advanced atherosclerotic lesions in *Ldlr*^{-/-} *T39*^{-/-} mice. Mice on the *Ldlr*^{-/-} background were fed WTD for 20 weeks. Serum lipoprotein cholesterol levels in (a) male and (b) female mice after 2 weeks on WTD, $n = 5$ per genotype per gender. c, Proximal aorta atherosclerotic lesion area after 20 weeks on WTD, $n = 5$ animals per genotype per gender. Data

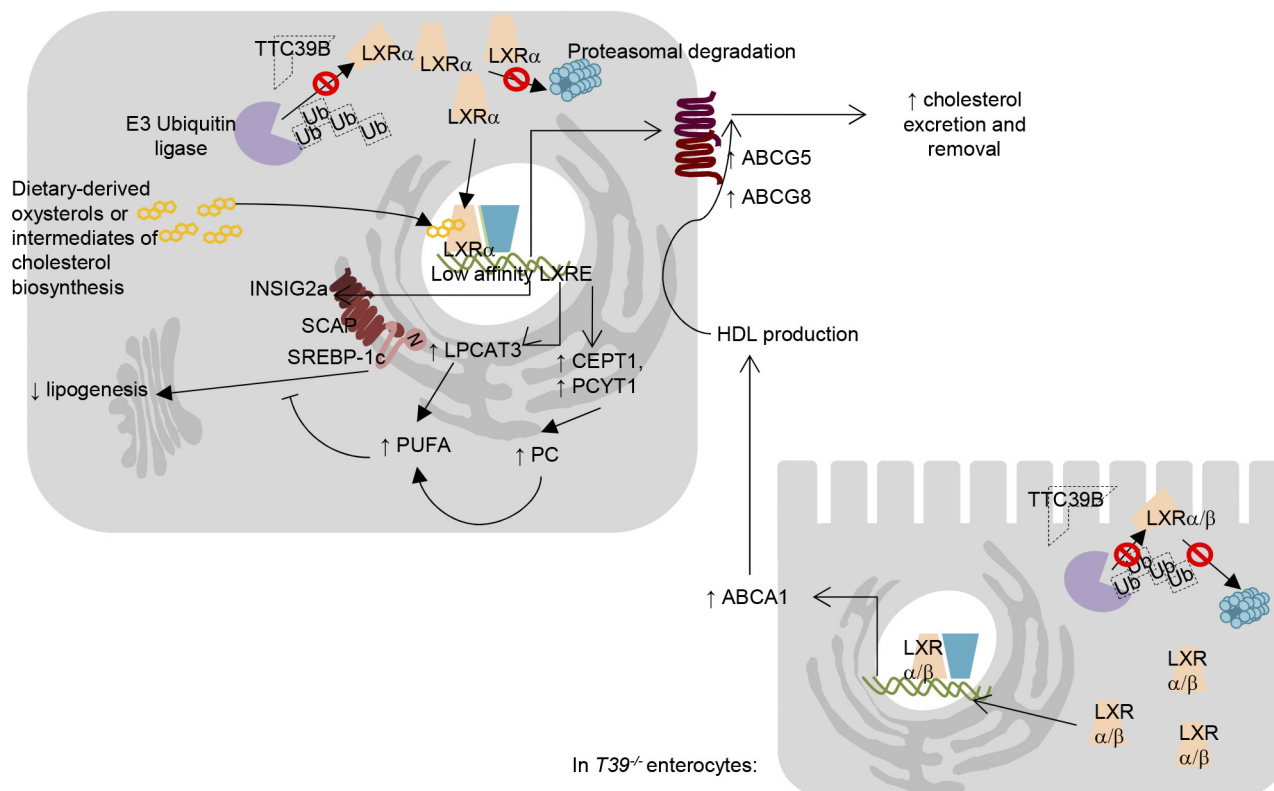
are represented as mean \pm s.e.m., $*P < 0.05$, $***P < 0.001$ by two-tailed Student's t -test. d, Lesion severity as graded by a blinded observer on six sections per animal, $n = 5$ per genotype per gender. Lesion severity is expressed as number of observations of each complexity category and the difference in the categorical distribution of lesions between the two groups is indicated, $*P < 0.05$ based on a χ^2 -test.



Extended Data Figure 8 | Increased LXR α protein in T39-deficient hepatocytes has implications for LXR target gene expression.

a, Immunoblots of endogenous LXR α in the nuclear (Nuc) and cytoplasmic (Cyto) fractions with an anti-LXR α antibody (top) 2 and 18 h after hepatocyte isolation and treatment with 2 μ M GW3965. Histone H3 (bottom) and Hsp90 (middle) are shown as loading controls for nuclei and cytoplasm, respectively. Normalization of LXR α signal to the appropriate loading control is shown below the LXR α blot. The immunoblot is representative of three different sets. **b**, RXR occupancy over LXREs of WT and T39 knockout primary hepatocytes treated with GW3965, $n = 4$ WT and 5 T39^{-/-}. **c**, Immunoblot showing validation of immobilized GW3965-mediated pulldown of endogenous LXR α from liver lysates. **d**, Relationship between magnitude of LXR target gene induction in T39 knockout mice and LXR affinity to active regulatory elements in the liver. See Supplementary Fig. 1 for gel source data.

In $T39^{-/-}$ hepatocytes:



In $T39^{-/-}$ enterocytes:

Extended Data Figure 9 | LXR protein preservation in $T39$ -deficient gastrointestinal tissues raises HDL and protects from steatohepatitis.

In the absence of $T39$, LXR assembly into a multiprotein complex that conjugates it to ubiquitin moieties does not occur, and LXR is spared from proteasomal degradation. In enterocytes, the increase in LXR α/β protein upregulates *Abca1* mRNA expression and promotes HDL production. In the liver, LXR α protein increase leads to the induction of *Abcg5/8*, which decreases dietary cholesterol uptake and increases cholesterol excretion, leading to cholesterol lowering. LXR α -mediated *Insig2a* prevents SREBP-1 processing in the fasted state, while *Pcyt1a*, and *Cept1* induction increases

microsomal membrane phosphatidylcholine content that continues to inhibit SREBP-1 processing in postprandial state. LPCAT3 upregulation results in increased incorporation of PUFA into phospholipid species which also contribute to the decrease in SREBP-1 processing. The decrease in nuclear SREBP-1 prevents the induction of lipogenic genes such as *Fasn* and *Pnpla3*. Therefore, unlike the gene expression profile that arises from potent synthetic ligands, increasing endogenous LXR protein levels preferentially upregulates cholesterol removal pathways while inhibiting lipogenesis.

Ki-67 acts as a biological surfactant to disperse mitotic chromosomes

Sara Cuylen¹, Claudia Blaukopf¹, Antonio Z. Politi², Thomas Müller-Reichert³, Beate Neumann⁴, Ina Poser⁵, Jan Ellenberg², Anthony A. Hyman⁵ & Daniel W. Gerlich¹

Eukaryotic genomes are partitioned into chromosomes that form compact and spatially well-separated mechanical bodies during mitosis^{1–3}. This enables chromosomes to move independently of each other for segregation of precisely one copy of the genome to each of the nascent daughter cells. Despite insights into the spatial organization of mitotic chromosomes⁴ and the discovery of proteins at the chromosome surface^{3,5,6}, the molecular and biophysical bases of mitotic chromosome structural individuality have remained unclear. Here we report that the proliferation marker protein Ki-67 (encoded by the *MKI67* gene), a component of the mitotic chromosome periphery, prevents chromosomes from collapsing into a single chromatin mass after nuclear envelope disassembly, thus enabling independent chromosome motility and efficient interactions with the mitotic spindle. The chromosome separation function of human Ki-67 is not confined within a specific protein domain, but correlates with size and net charge of truncation mutants that apparently lack secondary structure. This suggests that Ki-67 forms a steric and electrostatic charge barrier, similar to surface-active agents (surfactants) that disperse particles or phase-separated liquid droplets in solvents. Fluorescence correlation spectroscopy showed a high surface density of Ki-67 and dual-colour labelling of both protein termini revealed an extended molecular conformation, indicating brush-like arrangements that are characteristic of polymeric surfactants. Our study thus elucidates a biomechanical role of the mitotic chromosome periphery in mammalian cells and suggests that natural proteins can function as surfactants in intracellular compartmentalization.

To identify molecular factors that contribute to spatial separation of mitotic chromosomes we used an automated live-cell imaging pipeline. We visualized chromosome morphologies in HeLa cells stably expressing histone 2B (H2B) fused to a fluorescence resonance energy transfer (FRET) biosensor, which probes phosphorylation by the kinase Aurora B and thereby discriminates mitotic from interphase cells⁷. Addition of nocodazole excluded the effect of mitotic spindle perturbations. In this assay, we expected that depletion of any protein essential for chromosome separation would induce clusters of mitotic chromosomes, in contrast to the scattered chromosome distribution observed in control cells (Fig. 1a). We hence quantified the area of segmented chromosomes of live mitotic cells to detect clustering phenotypes (Fig. 1b).

Screening a small interfering RNA (siRNA) library targeting 1,295 candidate genes, including a comprehensive list of factors required for mitosis^{8,9} and components of the chromosome periphery⁶ (Supplementary Table 1), revealed a single hit with three different siRNAs causing a chromosome clustering phenotype: *MKI67*, the gene encoding the cell proliferation marker protein Ki-67 (Fig. 1c and Extended Data Fig. 1a, b). We validated on-target specificity of the RNAi phenotype by Cas9 nickase-mediated synonymous mutations

of the target region of one siRNA in all endogenous alleles of *MKI67* (Fig. 1d, e and Extended Data Fig. 1c–g).

Ki-67 is widely used as a proliferation marker in basic research and cancer prognosis^{10,11}, yet its molecular mechanism of action is unknown. There are some clues from the localization of Ki-67 to the chromosome surface from prophase until telophase^{12,13}, its requirement to target several other components of the mitotic chromosome periphery domain^{14,15}, and its interaction with the kinesin HKLP2 (ref. 16). Although Ki-67 does not seem to contribute to the internal structure of mitotic chromosomes¹⁴, its depletion causes nucleolar reassembly defects during mitotic exit¹⁴ and misorganized interphase heterochromatin¹⁵. Given that we had screened for chromosome separation regulators, we re-examined a potential role of Ki-67 in mitotic chromosome individualization.

We imaged mitosis in HeLa, non-cancer hTERT-RPE1 cells, and mouse embryonic stem cells depleted of Ki-67. Although metaphase plates in non-depleted control cells contained spatially separate chromosomes, they appeared as a single contiguous mass of chromatin in cells depleted of Ki-67 (Fig. 2a and Extended Data Fig. 2a–e). Time-lapse microscopy revealed that during prophase, when most chromosomes attach to the nuclear envelope¹, chromosomes condensed into separate bodies in Ki-67-depleted cells. However, soon after nuclear envelope breakdown (NEBD) chromosomes merged into a single coherent mass of chromatin (Fig. 2b and Extended Data Fig. 2f, g). Furthermore, in wild-type cells, Ki-67 relocalized from the nucleolus to chromosome arms only during very late prophase stages, after chromosome condensation had initiated (Extended Data Fig. 2h, i), and the internal structure of mitotic chromosomes appeared unaffected by Ki-67 depletion (Extended Data Fig. 3)¹⁴. Hence, Ki-67 is not required for the initial chromosome individualization and condensation during prophase, but for the maintenance of spatial separation after nuclear envelope breakdown.

The unstructured morphology of metaphase plates in cells depleted of Ki-67 might result from increased adhesion between neighbouring chromosomes. To test this, we tracked kinetochores in Ki-67-depleted cells stably expressing fluorescently labelled centromere protein A (CENP-A) after mitotic entry in the presence of nocodazole (Fig. 2c, d). In control cells, chromosomes moved extensively by free diffusion. Mitotic chromosomes of Ki-67-depleted cells were, in contrast, almost as immobile as interphase chromosomes in control cells (Fig. 2e, Extended Data Fig. 4a–c and Supplementary Video 1). Cell and nuclear size appeared normal and the nuclear envelope disassembled properly in Ki-67-depleted cells (Extended Data Fig. 4d–h and Supplementary Video 2). This suggests that mitotic chromosomes in Ki-67-depleted cells immobilize by increased adhesion rather than by spatial confinement.

A loss of spatial separation between mitotic chromosomes might impair spindle assembly and chromosome congression to the

¹Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna Biocenter (VBC), Dr. Bohr-Gasse 3, 1030 Vienna, Austria. ²Cell Biology and Biophysics Unit, European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, 69117 Heidelberg, Germany. ³Medical Faculty Carl Gustav Carus, Experimental Center, Technische Universität Dresden, Fetscherstrasse 74, 01307 Dresden, Germany. ⁴Advanced Light Microscopy Facility, European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, 69117 Heidelberg, Germany. ⁵Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, 01307 Dresden, Germany.

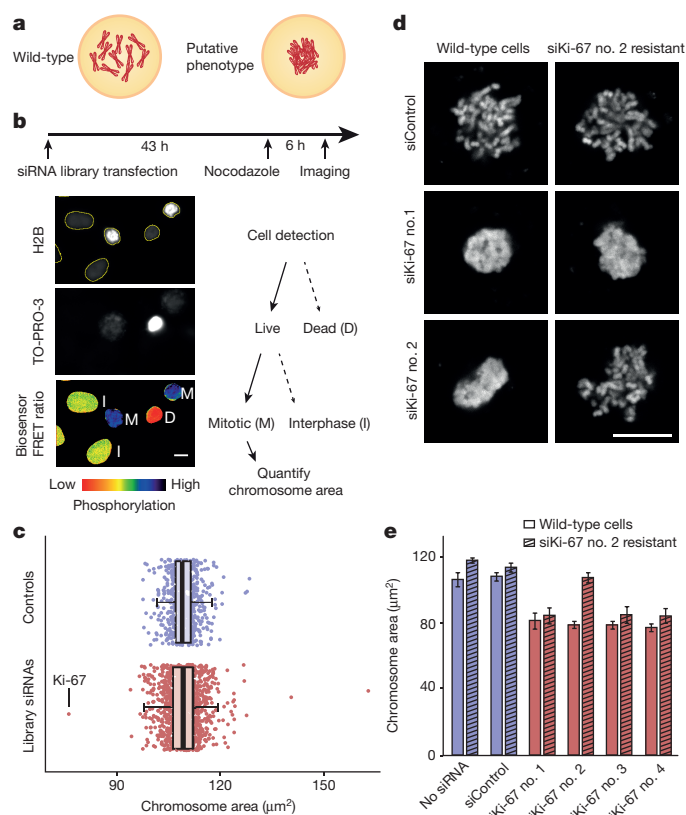


Figure 1 | RNAi screen for mitotic chromosome surface adhesion regulators. **a–c**, RNAi screen targeting 1,295 genes. **a**, Expected chromosome phenotypes in mitotic cells with depolymerized spindles. **b**, Experimental design to detect chromosome clustering in live mitotic HeLa cells based on the viability marker TO-PRO-3 and a FRET biosensor for mitotic phosphorylation. **c**, Individual data points correspond to the median chromosome area of all live mitotic cells per target gene, based on 2 or 3 different siRNAs. Mean, quartiles, and $1.5 \times$ interquartile range are indicated. **d**, **e**, Rescue of RNAi phenotype. **d**, Wild-type HeLa cells or a strain mutated in siKi-67 no. 2 target site, were transfected as indicated and imaged live in presence of nocodazole ($n = 9–10$ images per condition). **e**, Quantification of chromosome clustering as in **b**. The median size of the chromosome area was determined in 3 experiments (error bars indicate mean \pm s.d., $n > 71$ cells per condition and experiment, for exact sample numbers see Methods). Scale bars, $10 \mu\text{m}$.

metaphase plate. Ki-67 depletion indeed substantially delayed progression from nuclear envelope disassembly until anaphase onset (Extended Data Fig. 5a). Under these conditions, however, chromosomes were still separated at the onset of spindle assembly through their preceding tether to the nuclear envelope (Fig. 2b, 0 min)¹⁷, which might enable microtubule access to kinetochores before chromosomes coalesce. To investigate more specifically how chromosome clustering affects chromosome congression, we depolymerized the spindle by nocodazole during mitotic entry and then imaged spindle assembly after nocodazole washout. Most control cells rapidly formed metaphase plates and entered anaphase, whereas Ki-67-depleted cells consistently failed to assemble metaphase plates and almost never entered anaphase (Fig. 2f, g). Imaging cells stably expressing the microtubule plus-tip marker EB3–EGFP showed rapid growth of microtubules into regions between neighbouring chromosomes after nocodazole washout in control cells, but much less in Ki-67-depleted cells (Extended Data Fig. 5b, c). Hence, Ki-67-dependent chromosome separation is important for efficient access of spindle microtubules and progression to anaphase.

To address the mechanism by which Ki-67 establishes a non-adhesive chromosome surface, we aimed to map this function to

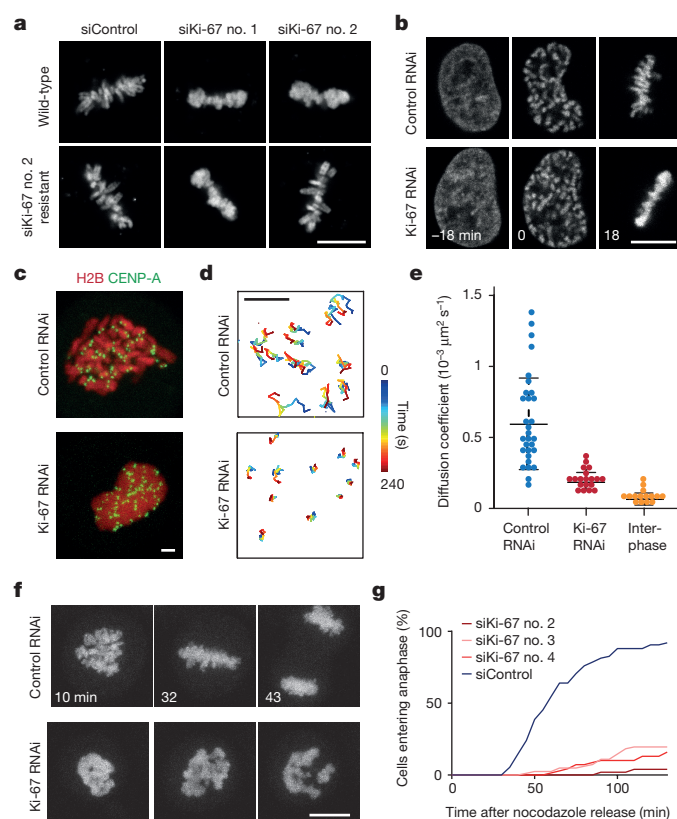


Figure 2 | Spatial separation of mitotic chromosomes by Ki-67 is important for chromosome motility. **a**, Live wild-type HeLa cells or the siKi-67 no. 2 resistant strain, transfected as indicated and stained with Hoechst. Quantified in Extended Data Fig. 2a. **b**, Time-lapse microscopy of HeLa cells expressing fluorescently-tagged H2B, transfected as indicated (NEBD, $t = 0$ min, $n = 16$ per condition). **c–e**, Kinetochore tracking. Live mitotic HeLa cells stably expressing H2B–mCherry/CENP-A–EGFP transfected with indicated siRNAs were imaged in presence of nocodazole (**c**) and kinetochores were tracked in time-lapse videos (**d**). **e**, Diffusion coefficients derived from mean square displacement (MSD) analyses. Error bars indicate mean \pm s.d. **f**, **g**, Live mitotic HeLa cells expressing fluorescently-tagged H2B imaged after nocodazole washout ($t = 0$ min) (**f**), were scored for anaphase entry (**g**) (cumulative frequency, 3 independent experiments with total sample numbers of $n = 75$ (siControl), $n = 51$ (siKi-67 no. 2), $n = 82$ (siKi-67 no. 3), $n = 69$ (siKi-67 no. 4)). Scale bars, $2 \mu\text{m}$ (**c**, **d**), all others $10 \mu\text{m}$.

specific domains. To test the function of Ki-67 truncation mutants (Fig. 3a), we generated a Ki-67 knockout cell line, which clustered mitotic chromosomes similar to the Ki-67 RNAi phenotype (Fig. 3b, Extended Data Fig. 6a–g and Supplementary Video 3). The observation that Ki-67 knockout cells survive despite the drastic chromosome coalescence phenotype might be explained by the initial separation of prophase chromosomes through tethers to the nuclear envelope¹⁷, which enables microtubule access to kinetochores before coalescence. However, the Ki-67 knockout cell line was sensitized to various stress conditions (Extended Data Fig. 6h), consistent with previous Ki-67 knockout studies^{15,18}.

The N terminus of Ki-67 contains a phosphopeptide-binding Forkhead-associated (FHA) domain¹⁹ and a protein phosphatase 1 (PP1)-binding site¹⁴. The central region consists of 16 tandem repeats²⁰ and the C-terminal part is enriched in leucine and arginine (LR) residue pairs²¹. As expected, expression of full-length Ki-67 restored mitotic chromosome individualization in Ki-67 knockout cells (no. 1, Fig. 3b). Ki-67 lacking the C-terminal LR domain does not bind to chromosomes¹² and it did not restore mitotic chromosome separation (no. 2, Fig. 3b) even if expressed at very high levels. In contrast, a construct lacking the entire N terminus, restored mitotic

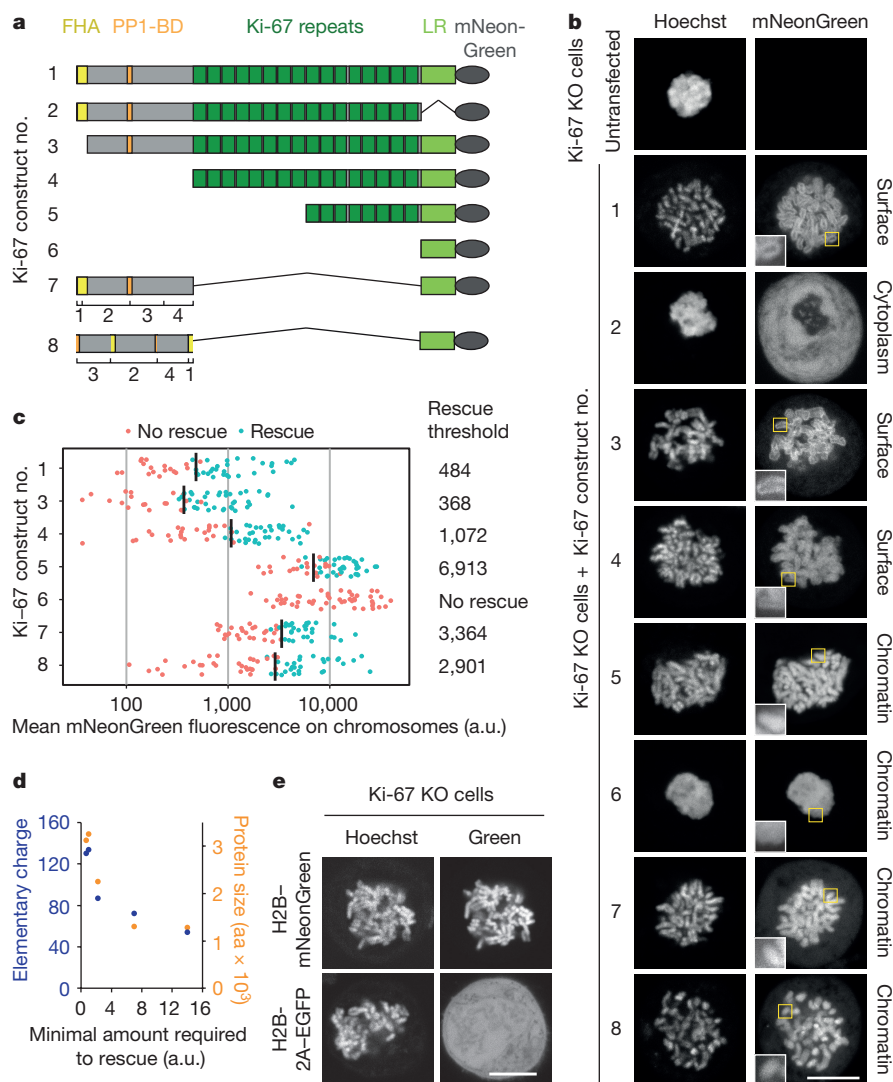


Figure 3 | The chromosome separation function of Ki-67 is not confined within a specific subdomain. a, Design of Ki-67 constructs. **b, c**, Ki-67 constructs were transfected in nocodazole-treated Ki-67 knockout HeLa cells (inserts are magnifications of yellow boxes) (**b**) and analysed for restoration of mitotic chromosome separation (**c**). Individual cells (dots) were classified by chromosome morphology (clustered: no rescue; dispersed: rescue) and plotted according to expression levels.

Black bars indicate threshold required for mitotic chromosome separation. a.u., arbitrary units. **d**, Correlation of Ki-67-truncation expression levels required for mitotic chromosome separation (as in **c**, relative to full length Ki-67) with their size (aa, amino acids) and predicted electrical charge at pH 7. **e**, Live Ki-67 knockout HeLa cells transfected with indicated constructs (quantified in Extended Data Fig. 7b). Scale bars, 10 μ m.

chromosome separation, indicating that the binding sites for the known Ki-67 interactors HKLP2 (ref. 16), NIFK²², and PP1 (ref. 14) are not relevant for this function (no. 4, Fig. 3b). A small Ki-67 fragment containing only half of the repeat domain and the chromosome-targeting LR domain still restored chromosome individualization, but the LR domain alone did not (no. 5 and no. 6, respectively, Fig. 3b). Unexpectedly, a complementary construct lacking the entire repeat domain also restored chromosome individualization (no. 7, Fig. 3b). Thus, any piece of the Ki-67 protein appears to support spatial separation as long as it is targeted to chromosomes by the LR-domain. To further test this, we designed a construct with randomly shuffled N-terminal fragments (no. 8, Fig. 3a). Remarkably, this construct also restored spatial separation of mitotic chromosomes in Ki-67 knockout cells (no. 8, Fig. 3b). Therefore the chromosome separating activity of Ki-67 is not confined within a specific protein region.

We wondered whether general physico-chemical features of the Ki-67 protein might provide clues about its mechanism. Ki-67 is very large (325 kDa and 360 kDa isoforms), has a very high net electrical charge (Extended Data Table 1), and is predicted to be mostly unfolded (Extended Data Fig. 7a). Furthermore, Ki-67 has an amphiphilic

structure, as the short C-terminal LR domain of Ki-67 has high attraction to chromatin, whereas its long N-terminal domain has high attraction to the cytoplasm and is excluded from chromatin (Fig. 3b, no. 6 and no. 2, respectively). An amphiphilic molecular structure combined with high electrical charge or long extensions (for example, long hydrocarbon chains) are typical features of surface-active agents (surfactants), which are chemical reagents that disperse particles or phase-separated liquid droplets²³. Through their localization at phase boundaries, surfactants stabilize dispersions and emulsions via steric hindrance and/or electrostatic repulsion²³. The striking similarity of Ki-67 to surfactants prompted us to further explore this function.

If Ki-67 acted by steric or electrostatic repulsion similar to surfactants, reducing its size or charge by truncating the peptide chain should decrease the efficiency of knockout phenotype rescue. We indeed observed a more than 14-fold difference in the amount of protein required to restore spatial separation of mitotic chromosomes between the minimal truncation version bearing only half of the repeats and the LR domain, and full-length Ki-67 protein (Fig. 3c). The phenotype rescue efficiency correlated with the protein size and predicted net charge of the constructs (Fig. 3d), suggesting that the

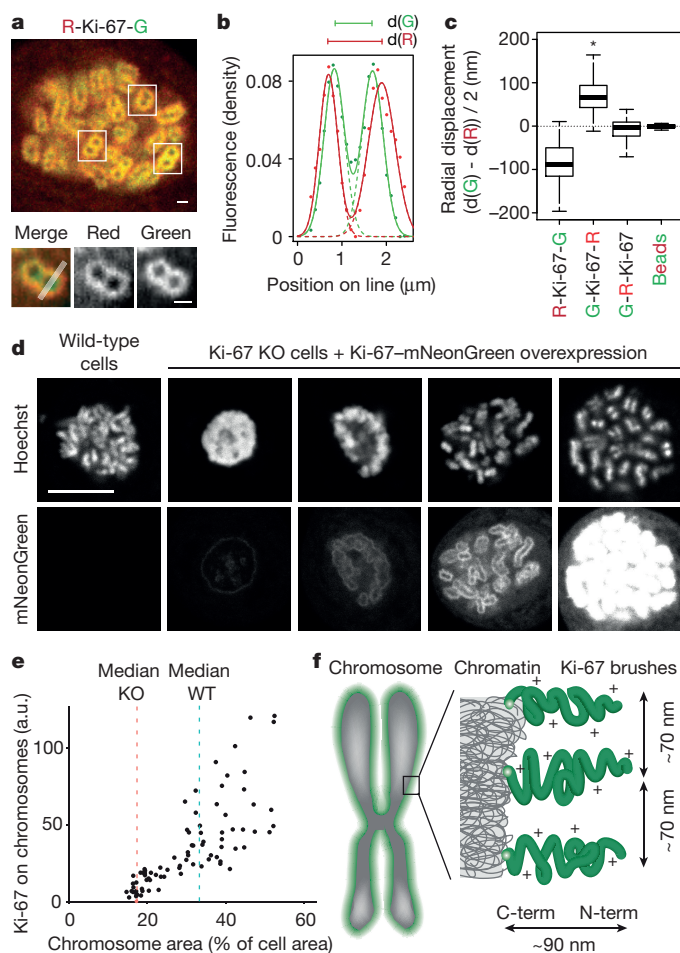


Figure 4 | Densely grafted Ki-67 molecules have an extended conformation perpendicular to the chromosome surface. **a, b**, Example of a live HeLa cell (from the analysis in **c**) expressing mCherry-Ki-67-EGFP. Chromosomes oriented perpendicular to the imaging plane (white boxes) were analysed by line profiles across one sister chromatid (bottom panels; white line) (**a**) and a sum of two Gaussian functions (solid line) was fitted to the line profile to measure peak-to-peak distance of the corresponding single Gaussians (dotted lines) (**b**). R, red (mCherry); G, green (EGFP). **c**, The average distance between the labelled Ki-67 polypeptide termini along the axis perpendicular to the chromosome surface was calculated for mCherry-Ki-67-EGFP ($n = 53$), EGFP-Ki-67-mCherry ($n = 52$), EGFP-mCherry-Ki-67 ($n = 44$) and multispectral fluorescent beads ($n = 34$). Boxes indicate median, quartiles and $1.5 \times$ interquartile range. **d, e**, Live nocodazole-treated Ki-67 knockout cells expressing different levels of Ki-67-mNeonGreen (**d**), were quantified with reference to chromosome area and Ki-67 levels on chromosomes (**e**) (each dot represents a cell, 4 independent experiments, dashed lines indicate the median of 39 wild-type and 40 Ki-67 knockout cells). **f**, Model of Ki-67-mediated mitotic chromosome repulsion. Scale bars, $1 \mu\text{m}$ (**a**), $10 \mu\text{m}$ (**d**).

size and overall electric charge might be important for the ability of Ki-67 to space mitotic chromosomes apart.

If electrical charge were a key mechanism of mitotic chromosome separation, then other highly positively charged DNA-binding proteins might be able to substitute for Ki-67. Strikingly, individually overexpressed core histones separated mitotic chromosomes in Ki-67 knockout cells — yet only at very high overexpression levels and not as far apart as Ki-67 (Fig. 3e and Extended Data Fig. 7b–h). This suggests that any positively charged chromosome-binding protein might separate mitotic chromosomes, yet Ki-67 apparently has additional properties relevant for spacing chromosomes further apart — for example, its large size and its enrichment at the mitotic chromosome surface.

To form an effective repulsive barrier, Ki-67 should cover a substantial fraction of the mitotic chromosome surface. To investigate this, we tagged all endogenous alleles of Ki-67 with enhanced GFP (EGFP) and measured the cytoplasmic concentration by fluorescence correlation spectroscopy (FCS) (Extended Data Fig. 8a–c). Using image segmentation of a reference DNA dye, we inferred that about 270,000 Ki-67 molecules bind to mitotic chromosomes, which corresponds to a surface density of about 210 Ki-67 molecules per μm^2 (Extended Data Fig. 8d–h), with an average spacing of 69 nm between Ki-67 molecules. Thus, Ki-67 is highly concentrated at the mitotic chromosome surface.

Polymeric surfactants adsorb at interfaces in a specific molecular orientation²³ and they form extended brush-like structures at high concentrations^{24,25}. Given the very C-terminal position of the chromatin-binding domain within Ki-67 and the exclusion of the remaining protein parts from chromatin regions (Fig. 3a, b), we probed the molecular extension of Ki-67 by attaching different fluorophores at each of its polypeptide ends (Extended Data Fig. 9a). Reference images of multispectral fluorescent beads showed that Gaussian fitting to fluorescence line profiles determines the relative localization of red and green fluorophores with an accuracy of $2.9 \pm 2.2 \text{ nm}$ (Extended Data Fig. 9b–d). We imaged live mitotic cells expressing mCherry-Ki-67-EGFP using confocal microscopy and determined the mean radial position of each fluorophore in perpendicularly sectioned chromosome arms (Fig. 4a, b). The N-terminal mCherry tag of Ki-67 localized $87.2 \pm 45.7 \text{ nm}$ outwards relative to the C-terminal EGFP tag, whereby a construct with inverted fluorophore positions and a control construct with both fluorophores attached to the N terminus of Ki-67 yielded consistent results (Fig. 4c and Extended Data Fig. 9e–h). Thus, Ki-67 has a very elongated conformation that orients perpendicular to the surface of mitotic chromosomes, consistent with a brush-like arrangement.

High grafting densities of Ki-67 molecules at the chromosome surface might increase the height of brush-like structures, as described for non-biological surface-attached polymers²⁴ (Extended Data Fig. 9i). Consistent with this, high overexpression levels of Ki-67 spaced mitotic chromosomes further apart than observed in wild-type cells (Fig. 4d, e). The range of Ki-67-mediated chromosome repulsion thus depends on molecular density.

Our study indicates that Ki-67 is required to maintain individual mitotic chromosomes dispersed in the cytoplasm after their release from the mechanically rigid nuclear envelope. Ki-67 might provide this function through a surfactant mechanism at the phase boundary between mitotic chromatin and the cytoplasm (Fig. 4f). As phase separation has emerged as an important principle underlying the formation of many other membrane-less cell organelles like nucleoli or centrosomes^{25–27}, it will be interesting to investigate whether and how natural protein surfactants might regulate other cellular phase boundaries.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 March; accepted 27 May 2016.

Published online 29 June 2016.

- Kschonsak, M. & Haering, C. H. Shaping mitotic chromosomes: From classical concepts to molecular mechanisms. *BioEssays* **37**, 755–766 (2015).
- Hirano, T. Condensin-based chromosome organization from bacteria to vertebrates. *Cell* **164**, 847–857 (2016).
- Ohta, S., Wood, L., Bukowski-Wills, J. C., Rappsilber, J. & Earnshaw, W. C. Building mitotic chromosomes. *Curr. Opin. Cell Biol.* **23**, 114–121 (2011).
- Naumova, N. *et al.* Organization of the mitotic chromosome. *Science* **342**, 948–953 (2013).
- Van Hooser, A. A., Yuh, P. & Heald, R. The perichromosomal layer. *Chromosoma* **114**, 377–388 (2005).
- Ohta, S. *et al.* The protein composition of mitotic chromosomes determined using multiclassifier combinatorial proteomics. *Cell* **142**, 810–821 (2010).
- Fuller, B. G. *et al.* Midzone activation of aurora B in anaphase produces an intracellular phosphorylation gradient. *Nature* **453**, 1132–1136 (2008).
- Neumann, B. *et al.* Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature* **464**, 721–727 (2010).

9. Hériché, J. K. *et al.* Integration of biological data by kernels on graph nodes allows prediction of new genes involved in mitotic chromosome condensation. *Mol. Biol. Cell* **25**, 2522–2536 (2014).
10. Whitfield, M. L., George, L. K., Grant, G. D. & Perou, C. M. Common markers of proliferation. *Nature Rev. Cancer* **6**, 99–106 (2006).
11. Yerushalmi, R., Woods, R., Ravdin, P. M., Hayes, M. M. & Gelmon, K. A. Ki67 in breast cancer: prognostic and predictive potential. *Lancet Oncol.* **11**, 174–183 (2010).
12. Saiwaki, T., Kotera, I., Sasaki, M., Takagi, M. & Yoneda, Y. *In vivo* dynamics and kinetics of pKi-67: transition from a mobile to an immobile form at the onset of anaphase. *Exp. Cell Res.* **308**, 123–134 (2005).
13. Verheijen, R. *et al.* Ki-67 detects a nuclear matrix-associated proliferation-related antigen. II. Localization in mitotic cells and association with chromosomes. *J. Cell Sci.* **92**, 531–540 (1989).
14. Booth, D. G. *et al.* Ki-67 is a PP1-interacting protein that organises the mitotic chromosome periphery. *eLife* **3**, e01641 (2014).
15. Sobacki, M. *et al.* The cell proliferation antigen Ki-67 organises heterochromatin. *eLife* **5**, e13722 (2016).
16. Vanneste, D., Takagi, M., Imamoto, N. & Vernos, I. The role of Hk1p2 in the stabilization and maintenance of spindle bipolarity. *Curr. Biol.* **19**, 1712–1717 (2009).
17. Beaudouin, J., Gerlich, D., Daigle, N., Eils, R. & Ellenberg, J. Nuclear envelope breakdown proceeds by microtubule-induced tearing of the lamina. *Cell* **108**, 83–96 (2002).
18. Cidado, J. *et al.* Ki-67 is required for maintenance of cancer stem cells but not cell proliferation. *Oncotarget* **7**, 6281–6293 (2016).
19. Hofmann, K. & Bucher, P. The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors. *Trends Biochem. Sci.* **20**, 347–349 (1995).
20. Schlüter, C. *et al.* The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiquitous nuclear protein with numerous repeated elements, representing a new kind of cell cycle-maintaining proteins. *J. Cell Biol.* **123**, 513–522 (1993).
21. Takagi, M., Matsuoka, Y., Kurihara, T. & Yoneda, Y. Chmadrin: a novel Ki-67 antigen-related perichromosomal protein possibly implicated in higher order chromatin structure. *J. Cell Sci.* **112**, 2463–2472 (1999).
22. Takagi, M., Sueishi, M., Saiwaki, T., Kametaka, A. & Yoneda, Y. A novel nucleolar protein, NIFK, interacts with the forkhead associated domain of Ki-67 antigen in mitosis. *J. Biol. Chem.* **276**, 25386–25391 (2001).
23. Rosen, M. J. & Kunjappu, J. T. *Surfactants and Interfacial Phenomena* 4th edn (John Wiley & Sons, 2012).
24. Milner, S. T. Polymer brushes. *Science* **251**, 905–914 (1991).
25. Hyman, A. A., Weber, C. A. & Jülicher, F. Liquid–liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **30**, 39–58 (2014).
26. Brangwynne, C. P., Mitchison, T. J. & Hyman, A. A. Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc. Natl Acad. Sci. USA* **108**, 4334–4339 (2011).
27. Zwicker, D., Decker, M., Jaensch, S., Hyman, A. A. & Jülicher, F. Centrosomes are autocatalytic droplets of pericentriolar material organized by centrioles. *Proc. Natl Acad. Sci. USA* **111**, E2636–E2645 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the IMBA/IMP BioOptics microscopy facility and Molecular Biology Service, J. Meissner, and M. J. Hossain for technical support, H. Liu and S. Tietscher for generation of plasmids, C. Haering, M. Samwer, W. H. Gerlich, and O. Wueseke for comments on the manuscript, Life Science Editors for editing assistance, and U. Kutay for LAP2 β -GFP/H2B-mRFP-expressing cells. D.W.G., A.A.H. and J.E. have received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 241548 (MitoSys), and A.Z.P., D.W.G. and J.E. under grant agreement no. 258068 (Systems Microscopy). D.G. has received funding from an ERC Starting Grant under agreement no. 281198 (DIVIMAGE), and from the Austrian Science Fund (FWF) project no. SFB F34-06 (Chromosome Dynamics). S.C. has received funding from a Human Frontier Science Program Long-Term Postdoctoral Fellowship and the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 330114 (IEF). T.M.R. was supported by Deutsche Forschungsgemeinschaft (DFG): SPP1384 'Mechanisms of Genome Haploidization' MU 1423/3-2 and grant MU 1423/8-1.

Author Contributions D.W.G. and S.C. conceived the project and designed experiments. B.N. and J.E. generated siRNA library transfection plates. I.P. and A.A.H. generated Ki-67-EGFP BAC cell pools. A.Z.P. and J.E. performed FCS measurements. D.W.G. and T.M.-R. performed electron microscopy experiments. S.C. performed all other experiments and C.B. assisted with the RNAi screen, chromosome spreads, cell line generation and cloning. D.W.G., S.C. and A.A.H. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.W.G. (daniel.gerlich@imba.oew.ac.at).

Reviewer Information *Nature* thanks T. Mitchison and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments. For outcome assessment most experiments were automatically analysed by machine learning or Fiji/MATLAB scripts to minimize potential human bias. When manual annotation was required, blinding precautions were made. In Fig. 3c and Extended Data Fig. 7b classification was performed based on the DNA channel alone without knowledge of the GFP signal. In Fig. 4c, line profiles were drawn without knowledge of which construct was transfected.

Cell lines and cell culture. All cell lines used in this study have been regularly tested negatively for mycoplasma contamination. Their sources and authentication is summarized in Supplementary Table 2. HeLa cell lines stably expressing fluorescent reporter proteins were generated from a HeLa Kyoto cell line as previously described²⁸. HeLa and hTERT-RPE1 cells were cultured in Dulbecco's modified Eagle medium (DMEM; Gibco) supplemented with 10% (v/v) fetal bovine serum (FBS; Gibco), 1% (v/v) penicillin-streptomycin (Sigma-Aldrich), 500 $\mu\text{g ml}^{-1}$ G418 (Gibco) and 0.5 $\mu\text{g ml}^{-1}$ puromycin (Calbiochem). CCE mouse embryonic stem cells were grown in gelatin-coated dishes in DMEM supplemented with 10% (v/v) FBS, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, 2 mM L-Glutamine, 0.05 mM β -mercaptoethanol, 1% penicillin-streptomycin, and 10^6 U ml^{-1} leukaemia inhibitory factor. To visualize chromatin cell lines either stably expressed histone H2B fused to mCherry or mRFP (Fig. 2c–e and Extended Data Figs 2h, i, 3c–g, 4a–c, e, f, h), fused to the Aurora B FRET biosensor (Figs 1b, c, 2b, f, g and Extended Data Figs 2f, g, 5a) or were labelled with Hoechst/SiR-Hoechst as indicated. For mutation of the endogenous Ki-67 locus in HeLa wild-type cells CRISPR/Cas9 nickase strategy was applied as described in Extended Data Fig. 1c, d. sgRNAs were cloned into pSpCas9n(BB)-2A-GFP (pX461, Zhang laboratory)²⁹. A circular plasmid (pCR2.1, Life Technologies) carrying the mutation and 700 bp homology flanks on each site was used as a repair template for homologous recombination. The plasmid mix of guide RNA plasmids and the repair template was transfected into HeLa cells using X-tremeGENE 9 DNA transfection reagent (Roche). Then 2 days after transfection cells were sorted for the presence of Cas9 (GFP positive), and another 4 days later for the absence of Cas9 (GFP negative) using fluorescence-activated cell sorting (FACS), either in batch or into 96-well plates, respectively. For endogenous EGFP tagging of Ki-67 in wild-type HeLa cells, sgRNAs were cloned into pSpCas9n(BB) (pX335, Zhang laboratory)²⁹ and 1,000 bp homology flanks were used. GFP positive cells were FACS sorted 6 days after transfection into 96-well plates. For live-cell imaging, HeLa and RPE1 cell lines were grown either in 96-well plastic-bottom plates (μ clear; Greiner Bio-One), or on LabTek II chambered coverglass (Thermo Scientific). Live-cell imaging was performed in DMEM containing 10% (v/v) FBS and 1% (v/v) penicillin-streptomycin, but without phenol red and riboflavin to reduce autofluorescence²⁸. CCE mouse embryonic stem cells were grown in LabTek II chambered coverglass (Thermo Scientific) coated with 5 $\mu\text{g ml}^{-1}$ Laminin-511 in PBS (BioLamina) for 2–3 h at 37 °C. The fluorescence correlation spectroscopy measurements were performed in a CO₂-independent imaging medium (Gibco) supplemented with 20% FBS (v/v), 1 mM L-glutamine, 1 mM sodium pyruvate and 100 nM SiR-Hoechst.

Electron microscopy. For electron microscopy, HeLa cells were grown on sapphire discs and cryo-immobilized using an EMPACT2 + RTS high-pressure freezer (Leica Microsystems, Vienna, Austria). Specimens were frozen in 100 μm -deep membrane carriers (Leica) filled with growth medium containing 20% BSA (Sigma). Samples were freeze-substituted in acetone containing 1% osmium tetroxide and 0.1% uranyl acetate and thin-layer embedded in Epon-Araldite as described in ref. 30. Serial thin sections were collected on Formvar-coated copper slot grids and imaged in a TECNAI 12 electron microscope (FEI, The Netherlands) operated at 100 kV.

Western blotting. Cells were transfected with siRNAs in 6-well plates (Thermo Scientific) and 2–3-million cells per ml were lysed in 1 \times SDS loading buffer at the indicated time points after transfection. Protein samples were separated on NuPAGE Novex 3–8% Tris-Acetate protein gels (Life technologies) and transferred to a nitrocellulose membrane (Protran BA83, Sigma) by semidry blotting. Ki-67 was probed by monoclonal anti-Ki-67 SP6 antibody (Abcam, ab16667) and actin by monoclonal anti-actin clone C4 (Milipore, MAB1501). Either fluorescently labelled antibodies (IRDye 800CW, IRDye 680RD; Odyssey) were used and membranes scanned on an Odyssey IR imager (LI-COR) or horseradish peroxidase-conjugated secondary antibodies (Bio-Rad) were visualized using ECL Plus Western Blotting Substrate (Thermo Scientific) and films.

Plasmid and siRNA transfection. For transient or stable expression of fluorescently tagged marker proteins, the genes were subcloned into IRESpuo2 vectors that allow expression of resistance genes and tagged proteins from a single transcript (Supplementary Table 3). Plasmids were transfected into HeLa cells using

X-tremeGENE 9 DNA transfection reagent (Roche) or FuGene 6 transfection reagent (Promega) according to the manufacturer's instructions and imaged 48 h post transfection. siRNAs (Supplementary Table 4) were delivered with Lipofectamine RNAiMax (Invitrogen) at a final concentration of 10 nM according to the manufacturer's instructions. XWneg9 and Scrambled were used as non-targeting siRNA controls and 4 different siRNAs against Ki-67 were used (see below). Ki-67 no. 1 corresponds to Ki-5 in ref. 14. Ki-67 siRNA no. 1 was used in Extended Data Figs 2b, c, e and 3c–h. Ki-67 siRNA no. 2 was used in Fig. 2b, Extended Data Figs 2f, g, 3i and 4h, Ki-67 siRNA no. 4 was used in Fig. 2c–e and Extended Data Figs 3a, b and 5b, c. All 4 Ki-67 siRNAs caused a similar chromosome coalescence phenotype as in Figs 1d and 2a. A nuclear shape phenotype as described in ref. 14 was only observed with Ki-67 siRNA no. 1, but not with the other siRNAs that depleted Ki-67 protein with similar efficiency (see Extended Data Figs 4d–g and 6f, g). All siRNAs were obtained from Life Technologies as Silencer Select reagents. Imaging was performed 48–72 h post-siRNA transfection as western blotting (Extended Data Figs 2b and 4d) indicated protein depletion to background levels.

Inhibitors and stains. Nocodazole (Sigma) was used at a final concentration of 100 ng ml^{−1} for 1–3 h to arrest cells in prometaphase. For washout experiments (Fig. 2f, g and Extended Data Fig. 5b, c), Labtek wells were washed 4 \times with 500 μl imaging medium without nocodazole before they were released into imaging medium. For acute addition during time-lapse imaging, 400 ng ml^{−1} was used (Extended Data Fig. 3g). Hoechst 33342 (Sigma) was used at a final concentration of 0.2 $\mu\text{g ml}^{-1}$. SiR-Hoechst³¹ at concentrations of 100–200 nM.

Ki-67 truncations. The long isoform of Ki-67 was cloned from HeLa cDNA into an IRESpuo2 vector under a CMV promoter and fused to mNeonGreen at the C terminus. Truncation mutants were generated by PCR amplification, enzymatic digest and ligation from the full length construct (3,256 amino acids) and cloned into the same plasmid. For truncation construct no. 2, amino acids 2929–3256 were removed; for construct no. 3, amino acids 1–134; for construct no. 4, amino acids 1–1002; for construct no. 5, amino acids 1–1970; for construct no. 6, amino acids 1–2930; and for construct no. 7, amino acids 995–2945. For construct no. 8, the N-terminal region from amino acids 1–994 was divided into 4 pieces and the order of the fragments was rearranged. The final construct contained the Ki-67 fragments in the following order: amino acids 507–726, amino acids 52–506, amino acids 727–994, amino acids 1–51 from N to C terminus.

RNAi screen. A total of 1,295 genes were targeted by either two or three siRNAs. This target gene list included the MitoCheck genome-wide RNAi screen validation data set with 1,128 genes⁸, 100 predicted chromosome condensation factors⁹ and 67 known and predicted chromosome periphery proteins (43 based on literature search, and additionally 24 based on Gene Ontology term and domain analysis in the data set presented in ref. 6). All siRNAs had been mapped against the 2013 human genome (ENSEMBL V70) to ensure unique target specificity. siRNAs were delivered using solid-phase reverse transfection³² in 384-well imaging plates (Falcon). Cells were seeded with 1 μM TO-PRO-3 Iodide (Life Technologies) on the screening plates using a Multidrop Reagent Dispenser (Thermo Scientific). 43 h after seeding, nocodazole was added with the same device to a final concentration of 100 ng ml^{−1} and 6 h later plates were imaged on a Molecular Devices ImageXpressMicro XL screening microscope (see below) using a $\times 20$, 0.75 NA, S Fluor dry objective (Nikon) and acquiring 4 positions with 520 $\mu\text{m} \times 520 \mu\text{m}$ in each well. To compensate for inhomogeneous illumination, all images were flatfield corrected with the Metamorph software (Molecular Devices) using background images acquired in empty wells. Automated image analysis was performed using the in-house-developed CellCognition software³³ and nuclei and sets of mitotic chromosomes were segmented by local adaptive thresholding. As a quality control for RNAi efficiency and specificity, supervised classification of cell morphologies was applied to the screening data and expected phenotypes⁸ were observed for all positive controls (siRNAs targeting INCENP, KIF11, PLK1, CDC20) and negative controls (empty, Scrambled siRNA, XWneg9 siRNA). For chromosome area measurements, only live mitotic cells were taken into account using a two-step gating strategy. First, mitotic cells were identified using the Aurora B FRET biosensor as depicted in Fig. 1b. Aurora B phosphorylation of the biosensor at the onset of mitosis causes a conformational change in the sensor that reduces FRET between a CFP for energy transfer (CyPet) donor and a YFP for energy transfer (YPet) acceptor⁷. To determine the range of FRET/YPet ratio of mitotic cells, nocodazole-arrested cells were identified by supervised classification using images from two untransfected wells of each plate. The 2.5th and 97.25th quantile of the background-subtracted FRET/YPet intensity of all prometaphase cells was used as lower and upper classification border, respectively ($0.6 < \text{FRET/YPet} < 0.82$). To identify dead cells and exclude them from further analysis, a TO-PRO-3 intensity threshold was applied ($\text{TO-PRO-3} > 15$). From all live mitotic cells of each well, the median size of the segmented chromosome area was calculated and plotted either for each siRNA individually (Extended Data Fig. 1a) or as the mean of 2–3 siRNAs

(Fig. 1c). Wells with >20% apoptotic cells or <20 live mitotic cells were excluded from the final analysis, as the chromosome area could not be reliably quantified under these conditions.

Live-cell microscopy. Automated wide-field fluorescence microscopy (Fig. 1a–c, e) was performed on a Molecular Devices ImageXpressMicro XL screening microscope equipped with reflection-based laser autofocus and a $\times 20$, 0.75 NA, S Fluor dry objective (Nikon), controlled by in-house-developed Metamorph macros³³. Cells were maintained in a microscope stage incubator at 37 °C in a humidified atmosphere of 5% CO₂. To image the Aurora B FRET cell line, the following two filter sets were used: a FRET filter cube with an excitation filter 426–450 nm, emission filter 528.5–555.5 nm and a dichromatic mirror of 458 nm and a YFP filter cube with an excitation filter 488–512 nm, emission filter 528.5–555.5 nm, and a dichromatic mirror of 520 nm.

Confocal microscopy was performed on a customized Zeiss LSM780 microscope using a $\times 40$, 1.4 NA, Oil DIC Plan-Apochromat objective (Zeiss), controlled by ZEN 2011 software and an autofocus macro (AutofocusScreen, <http://www.ellenberg.embl.de/index.php/software>) provided by J. Ellenberg. The microscope was equipped with an incubation chamber (European Molecular Biology Laboratory (EMBL), Heidelberg, Germany), providing a humidified atmosphere at 37 °C with 5% CO₂.

Fast time-lapse imaging with the EB3-GFP cell line was performed on a spinning-disk confocal microscope (UltraView VoX, PerkinElmer) with a $\times 100$, 1.45 NA objective controlled by Volocity software and equipped with an incubation chamber (EMBL) for imaging at 37 °C with 5% CO₂.

For fluorescence recovery after photobleaching (FRAP) experiments, selected image regions were bleached using a laser intensity 600-fold higher than the laser intensity used for image acquisition, and the pixel dwell time was increased 20-fold above that used for image acquisition.

Immunostaining of mouse embryonic stem cells. Mouse embryonic stem cells were fixed in 3.7% formaldehyde in PBS for 10 min, washed twice with PBS containing 0.05% Tween20 and permeabilized with 0.5% Triton X-100 in PBS. Cells were blocked for 10 min with 10% FBS in PBS containing 0.05% Tween20 and stained for 2 h with a rabbit polyclonal anti-Ki-67 antibody (Abcam, ab66155, 1:300). After several washes cells were stained with anti-rabbit IgG Alexa Fluor 488 (Molecular Probes, 1:600). DNA was stained with 0.8 $\mu\text{g ml}^{-1}$ Hoechst 33342 (Sigma).

Mitotic chromosome spreads. HeLa cells treated with siRNA for 48 h were trypsinized and resuspended in 75 mM KCl for 16 min at 37 °C. Cells were then fixed by 3:1 ice-cold methanol:acetic acid for 15 min at 4 °C. After 2 washes with 3:1 ice-cold methanol:acetic acid cells were dropped on cleaned and pre-chilled glass slides from a height of 30 cm. Cells were dried on the slide and mounted in Vectashield mounting medium with 1.5 $\mu\text{g ml}^{-1}$ DAPI (Vector Laboratories).

Micrococcal nuclease hypersensitivity assay. HeLa cells treated with siRNA for 48 h were trypsinized and washed twice with PBS. 10 million cells per sample were resuspended in lysis buffer (25 mM Tris pH 7.5, 100 mM NaCl, 5 mM MgCl₂, 0.2% NP-40, 1 \times Complete EDTA-free protease inhibitor (Roche)) and lysed on ice by passing 20 times through a needle (0.4 mm). Chromatin fraction was pelleted by centrifugation, washed once with micrococcal nuclease (MNase) digestion buffer (10 mM Tris pH 7.4, 15 mM NaCl, 60 mM KCl, 1 mM CaCl₂), resuspended

in MNase digestion buffer and aliquoted. Aliquots were prewarmed to 37 °C for 2 min and subsequently incubated with threefold serial dilutions of micrococcal nuclease from 0.06 U ml⁻¹ to 45 U ml⁻¹ (Thermo Scientific) for 20 min at 37 °C in MNase digestion buffer. The reaction was stopped by addition of 25 mM EDTA. 1% SDS was added, the sample was diluted 1:4 with deionized water and 500 mM NaCl was added. DNA was phenol–chloroform extracted and run on a 1.5% agarose gel.

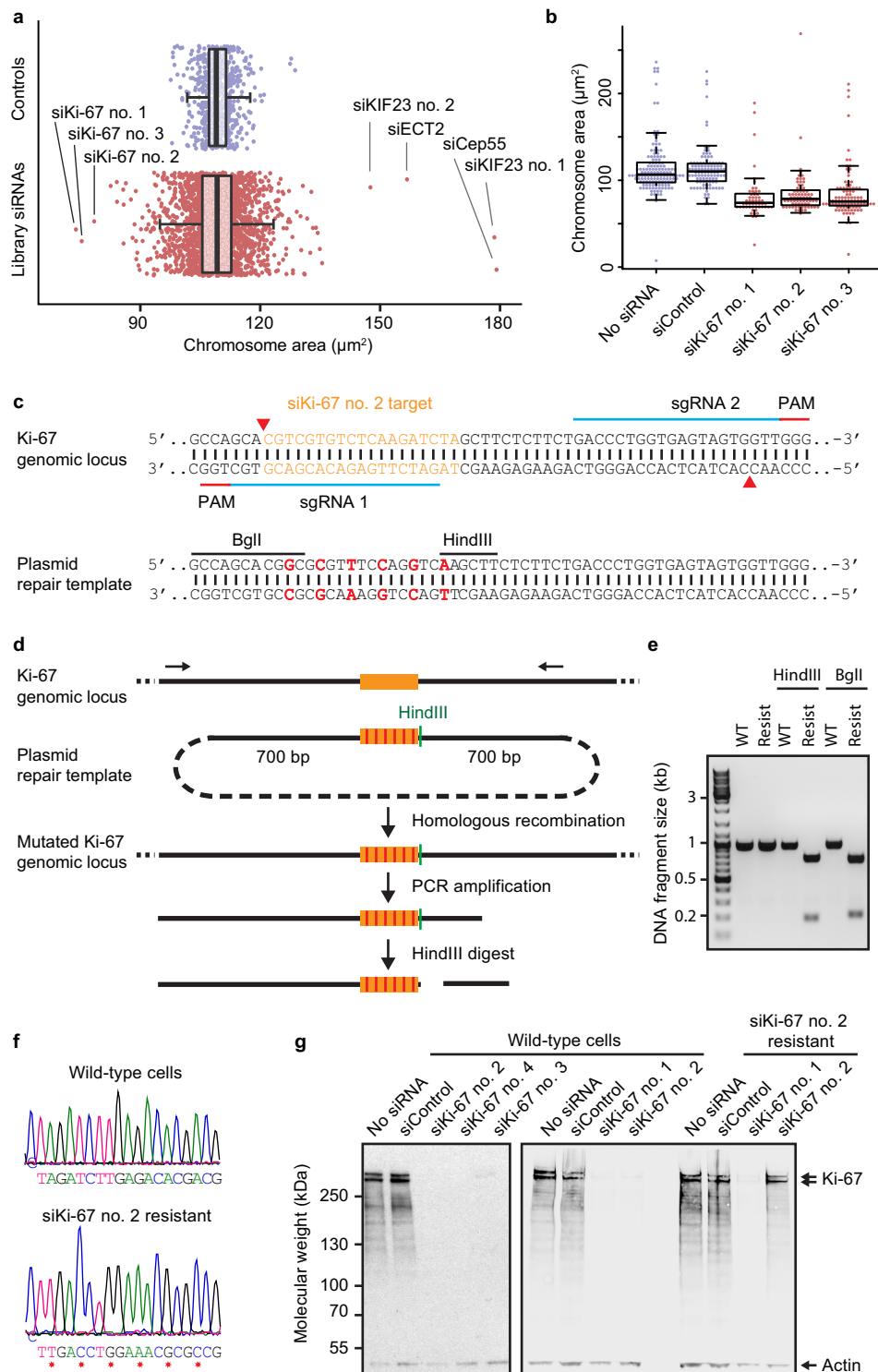
Colony formation assay. 200 wild-type HeLa or Ki-67 knockout cells were seeded into 6-well plates. After two days indicated drugs were added at increasing concentrations and plates were incubated for 10–17 days. Colonies were then fixed with 4% formaldehyde for 20 min, washed with water, stained for 20–30 min with Crystal Violet, washed with water several times and dried.

Secondary structure and charge predictions. Folded and charged regions within Ki-67 were visualized with FoldIndex (<http://bip.weizmann.ac.il/fldbin/findex>) and the EMBOSS charge prediction tool (<http://www.bioinformatics.nl/cgi-bin/emboss/charge>). A proteome-wide net charge analysis (Extended Data Table 1) was performed on the human proteome assembly GRCh38 using a customized R script based on the net charge calculation function of the 'seqinr' R package and pK values from EMBOSS.

Statistical analysis and sample numbers. All experiments were repeated several times and indicated experiment numbers always refer to biological replicates. Data were tested for normality and equal variances with Shapiro–Wilk and Levene's tests ($\alpha = 0.05$), respectively. The appropriate statistical test was chosen as follows: Unpaired normal distributed data were tested with a two-tailed *t*-test (in case of similar variances) or with a two-tailed *t*-test with Welch's correction (in case of different variances). Unpaired not normal distributed data were tested with two-tailed Mann–Whitney test (in case of similar variances) or with a two-tailed Kolmogorov–Smirnov test (in case of different variances). Paired not normal distributed data were tested with a Wilcoxon matched-pairs signed rank test.

Sample numbers of Fig. 1e: 3 experiments with the following sample numbers in wild-type cells: no siRNA ($n = 435$; 242; 327), siControl ($n = 856$; 472; 506), siKi-67 no. 1 ($n = 422$; 241; 201), siKi-67 no. 2 ($n = 480$; 318; 331), siKi-67 no. 3 ($n = 961$; 444; 466), siKi-67 no. 4 ($n = 859$; 560; 492). 3 experiments with the following sample numbers in siKi-67 no. 2 resistant cells: no siRNA ($n = 331$; 325; 235), siControl ($n = 932$; 291; 367), siKi-67 no. 1 ($n = 427$; 227; 71), siKi-67 no. 2 ($n = 498$; 223; 185), siKi-67 no. 3 ($n = 331$; 385; 298), siKi-67 no. 4 ($n = 847$; 223; 245).

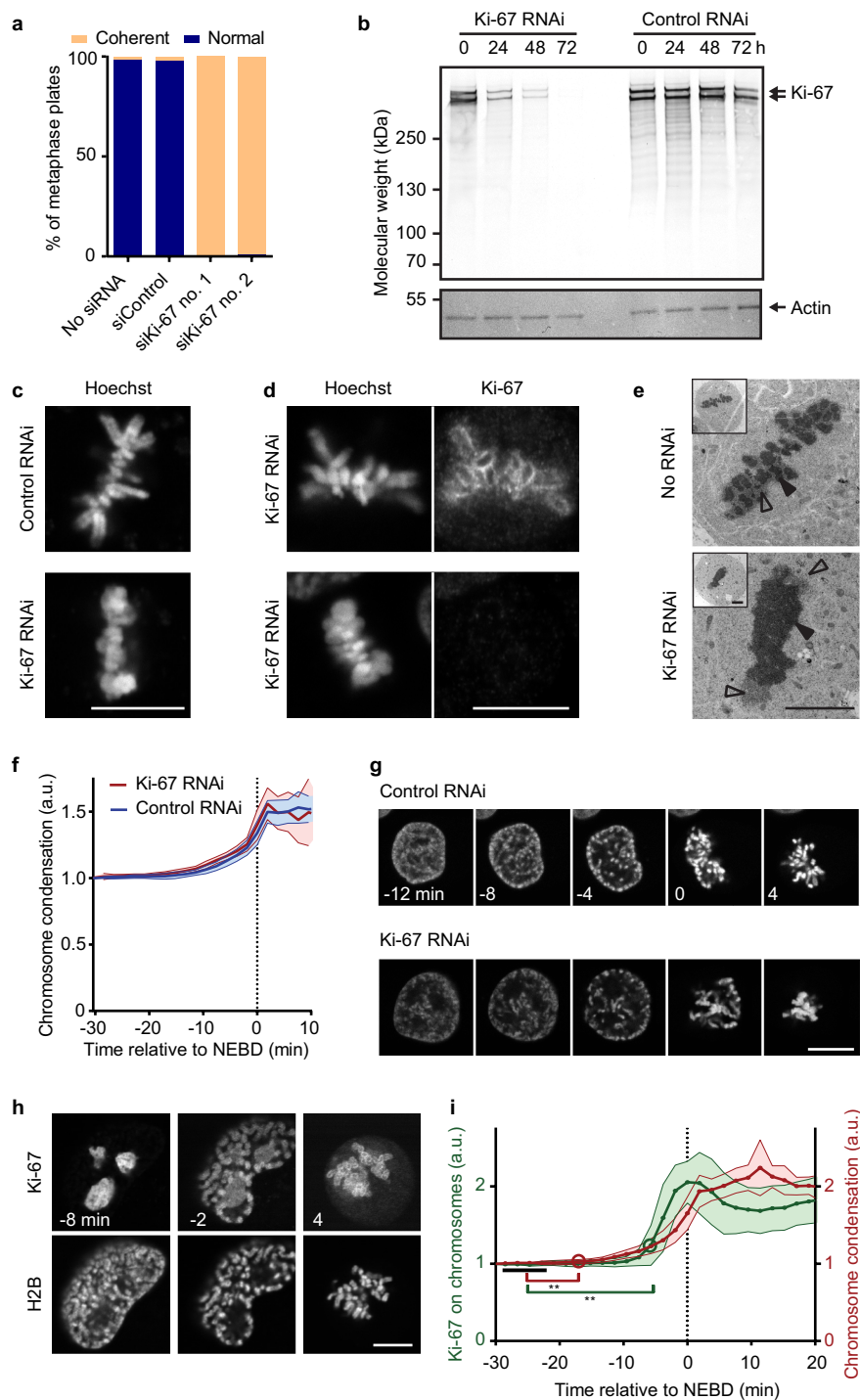
28. Schmitz, M. H. & Gerlich, D. W. Automated live microscopy to study mitotic gene function in fluorescent reporter cell lines. *Methods Mol. Biol.* **545**, 113–134 (2009).
29. Ran, F. A. *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380–1389 (2013).
30. Guizetti, J. *et al.* Cortical constriction during abscission involves helices of ESCRT-III-dependent filaments. *Science* **331**, 1616–1620 (2011).
31. Lukinavičius, G. *et al.* SiR-Hoechst is a far-red DNA stain for live-cell nanoscopy. *Nature Commun.* **6**, 8497 (2015).
32. Erfle, H. *et al.* Work flow for multiplexing siRNA assays by solid-phase reverse transfection in multiwell plates. *J. Biomol. Screen.* **13**, 575–580 (2008).
33. Held, M. *et al.* CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nature Methods* **7**, 747–754 (2010).



Extended Data Figure 1 | See next page for caption.

Extended Data Figure 1 | Generation of a Ki-67 siRNA no. 2 resistant HeLa cell line by homozygous mutation of the endogenous Ki-67 genomic loci. **a**, Quantification of chromosome area as in Fig. 1b, but displayed for each individual siRNA, showed that the three siRNAs causing strongest chromosome clustering all target Ki-67. Each data point corresponds to the median chromosome area of all live mitotic cells in a specific siRNA condition. Median, quartiles and $1.5 \times$ interquartile range of controls (siControls and untransfected) and the siRNA library are indicated. The top four siRNAs causing an increased chromosome area all target proteins involved in cytokinesis. Hence the chromosome area increase is likely to be a consequence of polyploidization by cytokinesis failure in preceding divisions. **b**, Chromosome areas of all individual live mitotic cells from the original RNAi screening data. Each data point corresponds to the chromosome area in a single cell. This reveals that the clustering phenotype is very penetrant within the cell population. **c**, CRISPR/Cas9 nickase strategy to mutate the siKi-67 no. 2 target site without changing the amino acid coding sequence. Red triangles indicate

DNA strand nicking sites. PAM, protospacer adjacent motif. sgRNA, single guide RNA. **d**, Schematic of genotyping strategy. A newly generated HindIII (or BglI) restriction site generated by CRISPR/Cas9 nickase as depicted in **c** was used to detect correctly mutated alleles. **e**, HindIII and BglI restriction fragments were detected by gel electrophoresis following the assay depicted in **d**, showing successful recombination of all three Ki-67 alleles present in HeLa cells. **f**, DNA sequencing chromatogram of the siKi-67 no. 2 target site of a wild-type and the CRISPR/Cas9-mutated cell line, respectively. Asterisks indicate mutated nucleotides. **g**, Western blot performed on whole cell lysates of wild-type or siKi-67 no. 2 resistant cell lines 48 h after the indicated siRNA transfections demonstrated that all siRNAs used in this study (siKi-67 nos 1–4) efficiently depleted Ki-67 in wild-type cells, and that the CRISPR/Cas9 mutated cell line was fully resistant against siKi-67 no. 2 but still sensitive to siKi-67 no. 1. The two bands labelled by anti Ki-67 antibody correspond to the two Ki-67 isoforms with predicted molecular masses of 320 and 359 kDa.

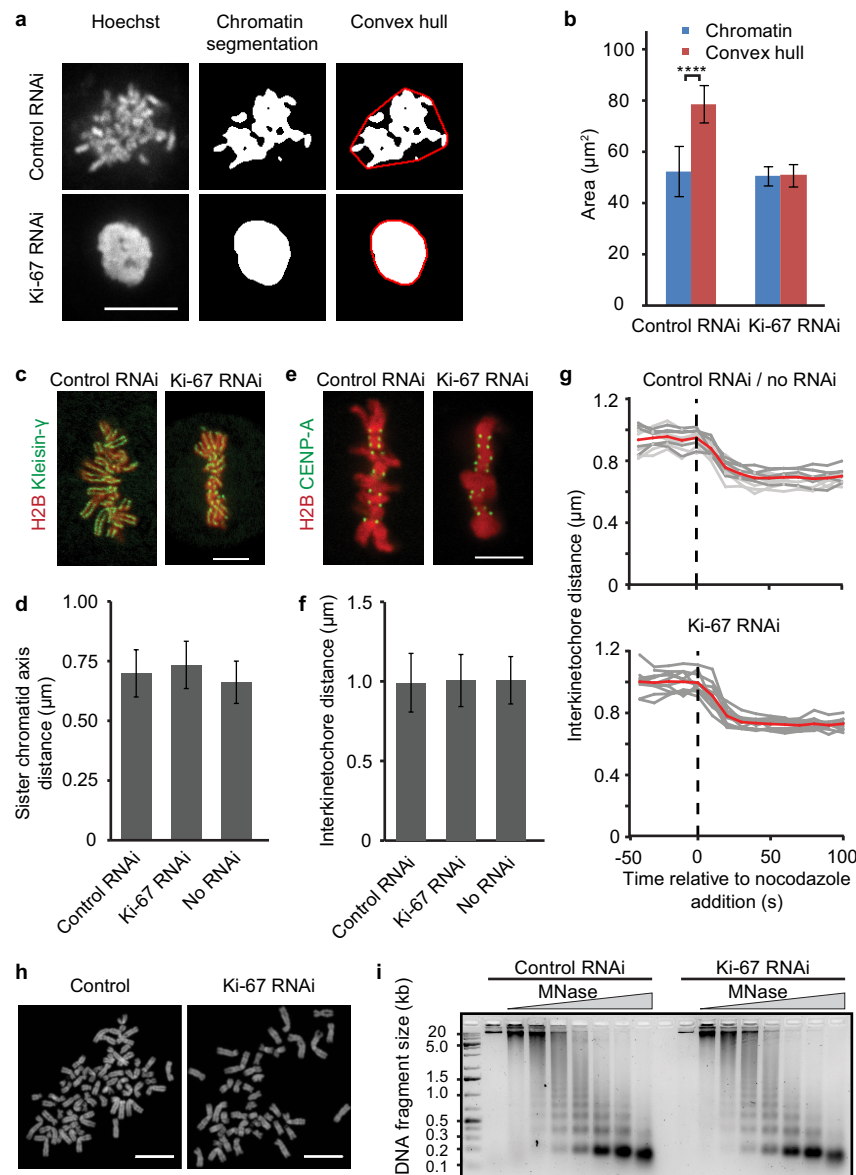


Extended Data Figure 2 | See next page for caption.

Extended Data Figure 2 | Ki-67 is not required for initial chromosome individualization and condensation but for maintenance of chromosome separation.

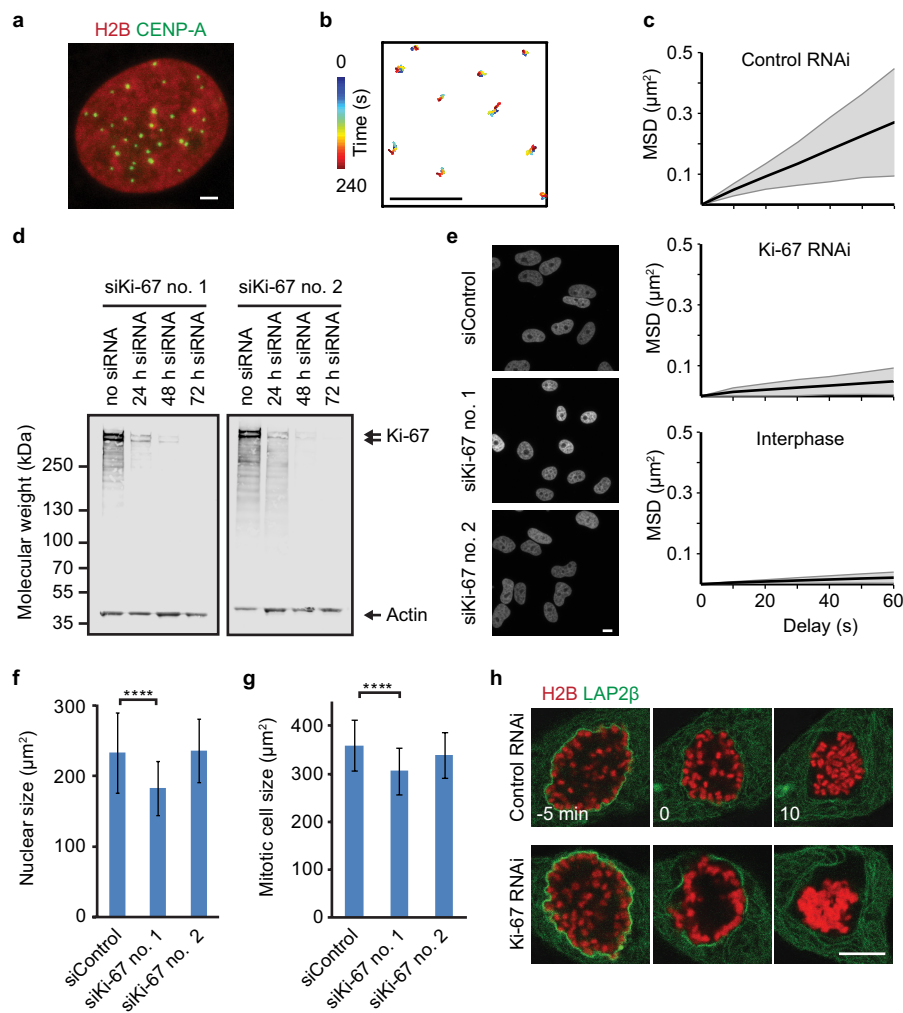
a, Quantification of phenotype penetrance in Fig. 2a. Live metaphase cells recorded by an automated imaging pipeline were classified as coherent or normal by supervised machine learning (3 independent experiments with total sample numbers of $n = 111$ (no siRNA), $n = 88$ (siControl), $n = 112$ (siKi-67 no. 1), $n = 131$ (siKi-67 no. 2)). **b**, Western blot analysis of HeLa wild-type cells, performed at indicated time points after siKi-67 no. 1 transfection, showed efficient depletion of Ki-67 at 48–72 h post-siRNA transfection. This time window was hence used for all further experiments. For uncropped gels, see Supplementary Fig. 1. **c**, Live hTERT-RPE1 cells stained with Hoechst were imaged 48 h after siRNA transfection. Control cells ($n = 21$) had spatially separate chromosomes, whereas Ki-67 siRNA-transfected cells ($n = 17$) had metaphase plates that appeared as a contiguous mass of chromatin. **d**, Mouse embryonic stem cells were fixed 48 h after siRNA transfection and stained using anti-Ki-67 antibody and Hoechst. Metaphase plates that lacked Ki-67 signal had chromosomes merged into an unstructured mass of chromatin (lower panel), whereas all cells with residual levels of

Ki-67, owing to incomplete RNAi-mediated protein depletion, had normal metaphase plate morphologies. Representative examples for $n = 20$ cells. **e**, Electron micrographs of HeLa cells transfected with siRNAs as indicated demonstrated that chromosomes appeared as a single contiguous mass. Representative images for $n = 10/10$ control and 9/11 for Ki-67 RNAi cells. Closed arrowheads mark dark chromatin areas; open arrowheads mark lighter grainy layer of chromosome periphery. **f**, Quantification of prophase chromosome condensation using pixel intensity standard deviation of the H2B-mCherry fluorescence as a measure for homogeneity of chromatin. Curves indicate mean \pm s.d. of 16 cells per condition. **g**, Representative example cells of the chromosome condensation assay in **f**. **h**, **i**, Ki-67 localization on chromosomes in live prophase HeLa cells expressing H2B-mCherry/Ki-67-EGFP (**h**) (NEBD, $t = 0$ min), was quantified (excluding nucleoli) during mitotic entry (**i**). A significant change ($P < 0.01$ by Student's t -test) in chromosome condensation (marked by red circle; compared to the initial 4 frames) occurs 11 min before a significant increase in Ki-67 localization (green circle). Curves indicate mean \pm s.d. of 13 cells. Scale bars, 5 μ m (**e**), all others 10 μ m.



Extended Data Figure 3 | Internal chromosome organization is not affected by Ki-67 depletion. **a, b**, Hoechst-stained chromosomes of control or Ki-67-depleted HeLa cells were segmented by thresholding, subsequently a convex hull was fitted around segmented chromosomes (red line) (**a**) and the segmented chromatin area and area of a convex hull fit were quantified (**b**). The calculated chromatin area of Ki-67 depleted cells was similar to control depleted cells, which demonstrated that Ki-67 depletion has no detectable effect on chromosome-internal compaction. Ki-67 depleted cells lacked inter-chromosomal space as the convex hull area was almost identical with the chromatin area ($n = 16$ for each siRNA, bars indicate mean \pm s.d., **** $P < 0.0001$ by paired Wilcoxon matched-pairs signed rank test). **c, d**, Live HeLa cell expressing H2B-mCherry/EGFP-Kleisin- γ were imaged 72 h after siKi-67 no. 1 or control siRNA transfection (**c**) and mean \pm s.d. of sister chromosome axis distance were quantified (**d**) ($n = 508$ from 3 experiments (control RNAi), $n = 380$ from 4 experiments (Ki-67 RNAi), $n = 398$ from 3 experiments (no RNAi)). **e, f**, Live HeLa cell stably expressing H2B-mCherry/CENP-A-EGFP were

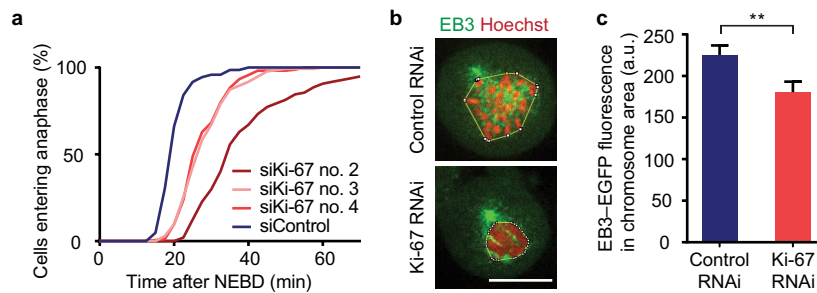
imaged ~ 60 h after siKi-67 no. 1 or control siRNA transfection (**e**) and mean \pm s.d. of interkinetochore distance were quantified (**f**) ($n = 875$ from 3 experiments (control RNAi), $n = 1,489$ from 5 experiments (Ki-67 RNAi), $n = 818$ from 3 experiments (no RNAi)). **g**, Centromere elasticity assay. Metaphase HeLa cells stably expressing H2B-mCherry/CENP-A-EGFP were imaged with 10 s time-lapse before and after nocodazole addition. The interkinetochore distance measurement over time in 12 control cells (control siRNA, light grey; untransfected cells, dark grey) and 10 cells transfected with siKi-67 no. 1 (grey lines) demonstrated that kinetochores were under tension in Ki-67 depleted cells. The red line indicates the mean. **h**, Chromosome spreads of untransfected or Ki-67 depleted HeLa cells appeared similar. Representative examples of 35 chromosome spreads from 2 independent experiments are depicted. **i**, Nuclei from control or Ki-67 depleted HeLa cells were treated with titrated amounts of MNase and DNA was subjected to 1.5% agarose electrophoresis. No detectable difference was seen in two independent experiments. Scale bars, 10 μm (**a, h**), 5 μm (**c, e**).



Extended Data Figure 4 | Chromosome clustering and motility decrease are not caused by spatial confinement.

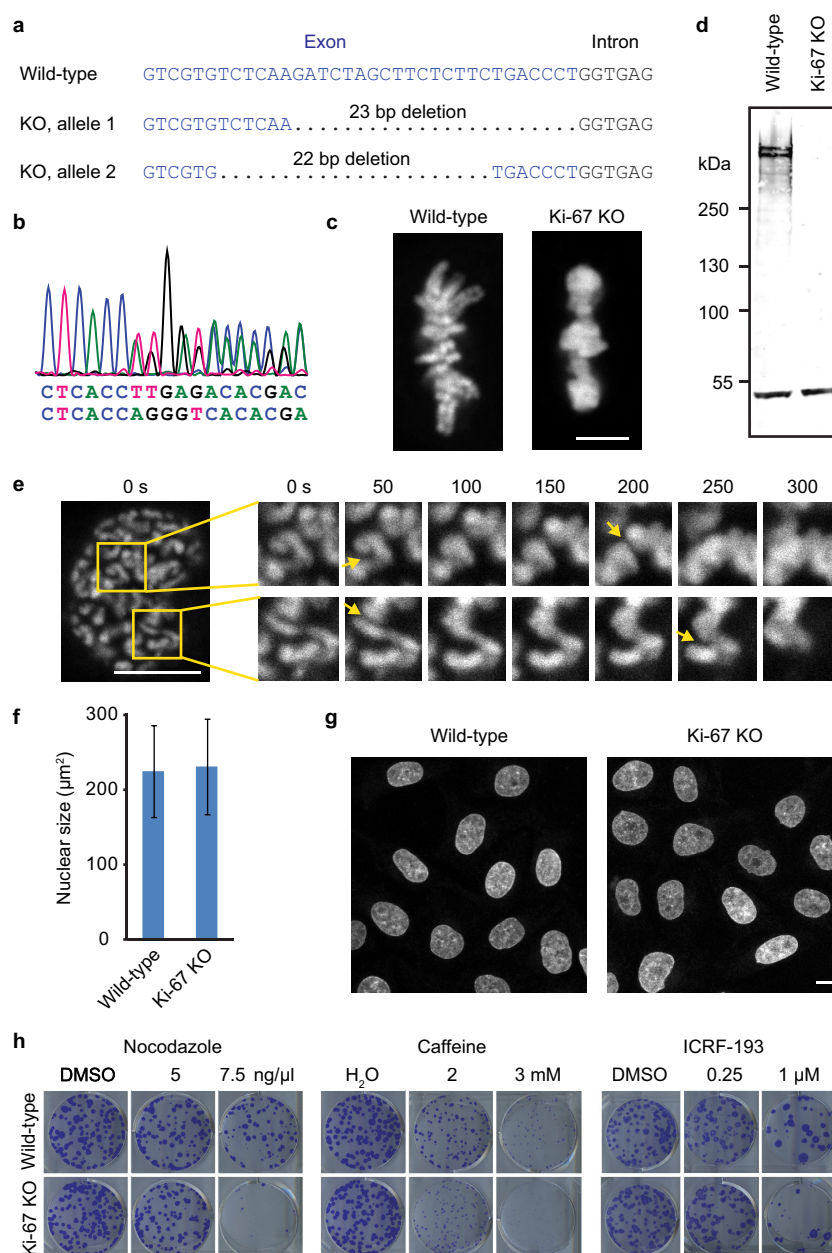
a, b, Representative example of kinetochore tracking performed in interphase (quantified in Fig. 2e). Live interphase HeLa cells stably expressing H2B-mCherry/CENP-A-EGFP were imaged in presence of nocodazole (**a**) and kinetochores were tracked in time-lapse videos (**b**). **c**, Representative MSD analyses of kinetochore tracks (CENP-A-EGFP) of mitotic HeLa cells (from Fig. 2c, d) and of the interphase cell in **a**. **d**, Western blot analysis of wild-type HeLa cells, performed at indicated time points after siKi-67 no. 1 or no. 2 transfection confirmed that both siRNAs deplete Ki-67 protein with similar efficiency. **e, f**, Interphase cells expressing H2B-mCherry 48 h after indicated siRNA transfections (**e**), were quantified with regard to nuclear size by automated segmentation of H2B-mCherry (**f**) (scale bars indicate mean \pm s.d., **** $P < 0.0001$ by Kolmogorov-Smirnov test, data from 25 random positions: $n = 303$ (siControl), $n = 303$ (siKi-67 no. 1), $n = 246$ (siKi-67 no. 2)). Only siKi-67 no. 1 caused a significant decrease

in nuclear size, while siKi-67 no. 2, which was used for kinetochore tracking in Fig. 2c-e, had no effect on nuclear size. As Ki-67 knockout cells also had a normal nuclear size (Extended Data Fig. 6f, g), we conclude that the effect of siKi-67 no. 1 on nuclear size is likely due to off-target protein depletion. **g**, Quantification of mitotic cell size of control or Ki-67 depleted cells demonstrated that similar to **f**, only siKi-67 no. 1 led to a significant decrease in cell size consistent with an off-target effect (error bars indicate mean \pm s.d., **** $P < 0.0001$ by Mann-Whitney test, total sample numbers: $n = 74$ (siControl), $n = 100$ (siKi-67 no. 1), $n = 95$ (siKi-67 no. 2)). As kinetochore tracking was based on siKi-67 no. 2, the reduced motility of mitotic chromosomes cannot be attributed to cell size changes. **h**, Live HeLa cells stably expressing H2B-mRFP/LAP2 β -EGFP were imaged 48 h after siRNA transfection as they entered mitosis in the presence of nocodazole ($n = 15$ per condition). Scale bars, 2 μm (**a, b**), 10 μm (**e, h**).



Extended Data Figure 5 | Ki-67 depletion causes a prolonged mitosis and impairs access of spindle microtubules to chromosomes. **a**, Mitotic progression determined by live-cell microscopy of HeLa cells expressing H2B-Aurora B FRET biosensor, after siRNA transfection as indicated, 3 independent experiments with total sample numbers of $n = 145$ (siControl), $n = 97$ (siKi-67 no. 2), $n = 134$ (siKi-67 no. 3), $n = 164$

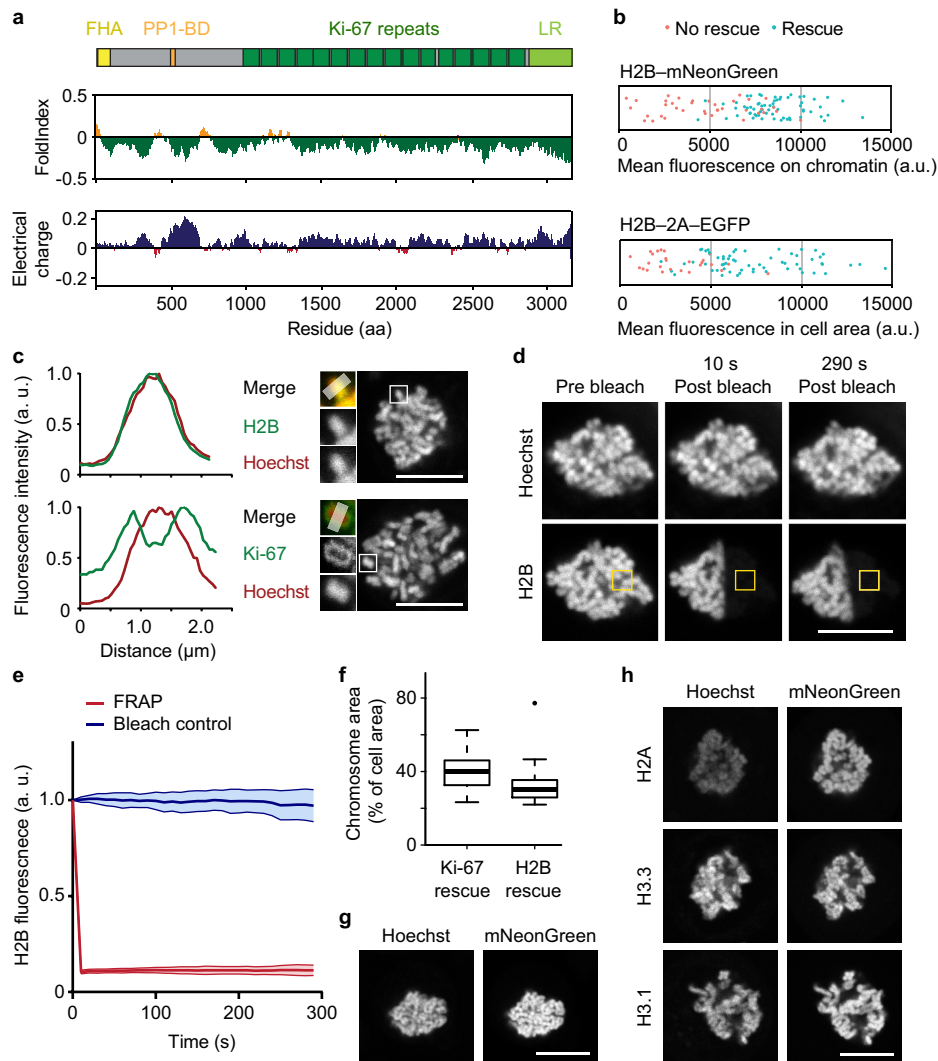
(siKi-67 no. 4). **b**, **c**, Live HeLa cells stably expressing EB3-EGFP were imaged 2 min after release from a 2 h nocodazole treatment, chromosome areas were identified (yellow lines) (**b**) and EB3-EGFP mean fluorescence was measured in chromosome areas (**c**) (error bars indicate mean \pm s.e.m. from 50 cells per condition, $**P < 0.01$ by Mann-Whitney test). Scale bar, $10\mu\text{m}$.



Extended Data Figure 6 | Generation of a Ki-67 knockout cell line.

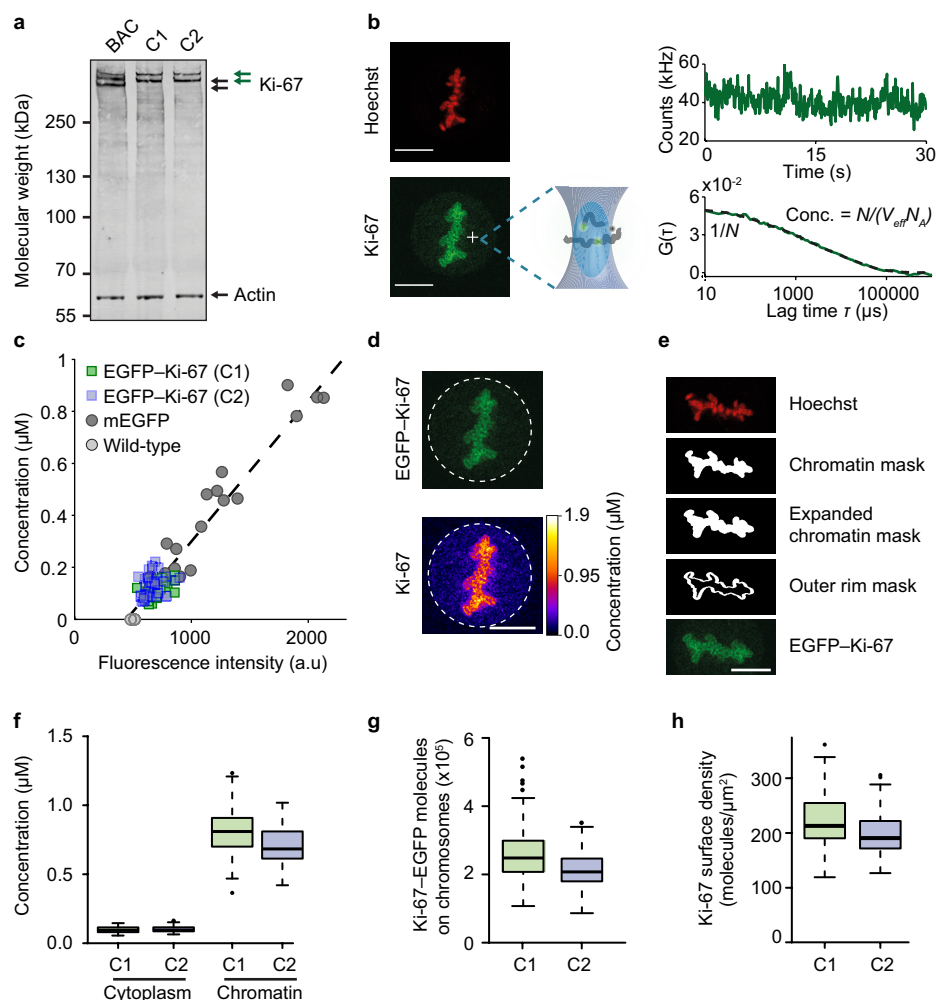
a, CRISPR/Cas9 was used to generate a HeLa cell line with indicated deletions on exon 5 of the Ki-67 allele. **b**, DNA sequencing chromatogram confirmed that no further alleles are present. **c**, Metaphase plates of live wild-type HeLa and Ki-67 knockout cells stained with Hoechst ($n = 30$ per cell line). **d**, Western blot performed on whole cell lysates of wild-type or Ki-67 knockout cells. The two high molecular weight bands labelled by anti-Ki-67 antibody in wild-type HeLa that corresponded to the two Ki-67 isoforms were undetectable for Ki-67 knockout cells. Actin was used as a loading control. **e**, Representative time-lapse image series of a Ki-67 knockout cell proceeding from prophase to prometaphase in the presence of nocodazole ($n = 12$, see Supplementary Video 3). Chromosomes are

labelled with SiR-Hoechst and two regions were selected to exemplify coalescence of chromosomes upon their close approach. Arrows mark regions just before their coalescence. **f**, Automated segmentation of SiR-Hoechst-labelled interphase nuclei of wild-type and Ki-67 knockout cells confirmed a normal nuclear size of Ki-67 knockout cells (mean \pm s.d. of 200 wild-type and 270 Ki-67 knockout cells). **g**, Representative example images of interphase wild-type or Ki-67 knockout cells stained with SiR-Hoechst of the quantification in **f**. **h**, The sensitivity of Ki-67 knockout cells to low dose nocodazole, caffeine or a topoisomerase II inhibitor (ICRF-193) was compared to wild-type by a colony formation assay. Representative images from two to three independent experiments are shown. Scale bar, 5 μm (**c**), 10 μm (**e**, **g**).



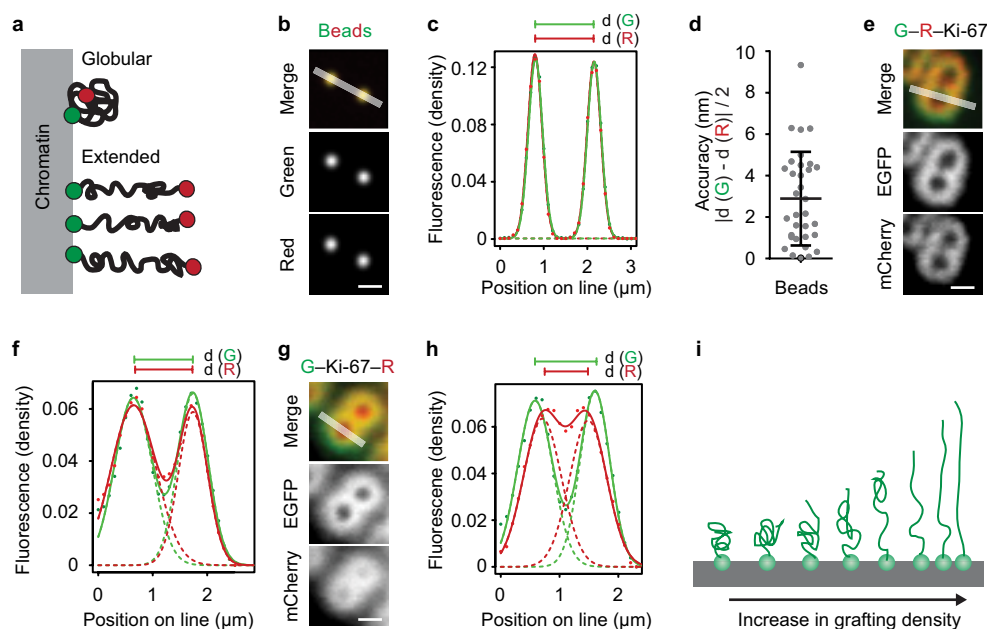
Extended Data Figure 7 | Ki-67 has little secondary structure, is highly positively charged, and its absence can be partly compensated by overexpression of core histones. **a**, Folding and charge prediction of full length Ki-67 based on FoldIndex and EMBOSS webtools using a sliding window of 100. Unfolded regions are depicted in green, folded regions in orange. Positive charge is marked in blue, negative charge in red. **b**, Quantification of overexpressed histone levels in individual cells related to the mitotic chromosome morphology phenotype, classified by visual inspection. Note that the mean fluorescence values are not comparable to Fig. 3c as different imaging settings had to be used. Cells are from 4–5 independent experiments. **c**, Radial localization of overexpressed H2B-mNeonGreen in live Ki-67 knockout cells ($n = 20$). Normalized fluorescence intensity along line profiles across a chromosome arm of live Ki-67 knockout cells transiently transfected with H2B-mNeonGreen (upper panel) or Ki-67-mNeonGreen (lower panel) indicate that overexpressed H2B binds to the surface as well as internal region within chromosomes. **d**, **e**, Stable association of H2B-mNeonGreen with mitotic chromosomes. **d**, Half of the mitotic chromosomes in Ki-67 knockout cells

highly overexpressing H2B-mNeonGreen were photobleached and the recovery of fluorescence was followed by time-lapse recording in an image region (yellow box). Representative example of the quantification in **e**. **e**, Curves indicate mean \pm s.d. of 20 photobleached and 19 unbleached control cells. **f**, Quantification of mitotic chromosome area relative to total cell area for cells shown in **b** and Fig. 3c. Boxes indicate median, quartiles and $1.5 \times$ interquartile range ($n = 30$ for Ki-67, $n = 64$ for H2B). **g**, Live Ki-67 knockout cell transiently transfected with a H2B-mNeonGreen and stained with SiR-Hoechst. Although 25 of 64 rescued cells displayed fully separated chromosomes (Fig. 3e), a large fraction of cells (39 of 64 rescued cells) showed detectable chromosome individualization at a lower extent compared to wild-type cells. Representative single z -section of the latter is shown. **h**, Live Ki-67 knockout cells transiently transfected with plasmids for expression of the indicated histone fused to mNeonGreen and stained with SiR-Hoechst. Representative single z -sections of 12–18 cells from 2–3 independent experiments are depicted. Although chromosome individualization was restored, chromosomes were not separated to the same extent as in wild-type cells. Scale bars, $10 \mu\text{m}$.



Extended Data Figure 8 | Fluorescence correlation spectroscopy of endogenous Ki-67 tagged with EGFP. **a**, Western blot performed on whole cell lysates of a HeLa cell line overexpressing GFP-tagged Ki-67 from a bacterial artificial chromosome (BAC), and of two different clones in which all endogenous Ki-67 alleles were N-terminally tagged with EGFP (C1, C2). The two lower bands labelled by anti-Ki-67 antibody in the BAC cell line (black arrows) correspond to the two wild-type Ki-67 isoforms. The EGFP-Ki-67 versions appear upshifted (green arrows) and the band of the small EGFP-tagged isoform overlaps with the band of the wild-type large isoform. **b–h**, FCS measurement from 3 independent experiments with 111 cells (C1) and 156 cells (C2). **b**, Imaging and analysis pipeline for FCS-calibrated imaging. Metaphase cells were identified based on Hoechst staining and imaged in 3D (only the central slice is shown). In the cytoplasm of the central slice an FCS measurement was performed. From the photon counts (right upper panel) the autocorrelation function (ACF) was computed (right lower panel). From the fit of the ACF to equation S1 in the Supplementary Information, the number of particles N in the focal volume was obtained. The concentration at the FCS point was computed by dividing N by the effective focal volume V_{eff} and the Avogadro's constant N_A . For details see Supplementary Methods. **c**, A calibration curve was obtained by plotting the concentration computed

from FCS against the fluorescence intensity in a 5×5 pixel region at the FCS measurement point (equation S2 in the Supplementary Information, dashed line). **d**, Fluorescence microscopy image of live HeLa cell with endogenous Ki-67, labelled by EGFP, scaled to absolute Ki-67 concentration as determined by FCS. **e**, Example segmentation of chromosomes and the chromosome surface used for quantification in f–h. Segmentation was performed in 3D, but for simplification only a single z -section is shown. A chromatin mask was obtained by segmentation of the Hoechst signal. To determine Ki-67 concentration on chromosomes, the mask was dilated to include Ki-67 signal at the outer chromosome surface (expanded chromatin mask). To calculate Ki-67 molecules on the outer chromosome surface, a rim around the chromatin mask was used (outer rim mask). See Supplementary Methods for details. **f–h**, Quantification of FCS measurements. Boxes indicate median, quartiles and $1.5 \times$ interquartile range. **f**, Mean Ki-67 concentration in the cytoplasm and on 3D-segmented chromosomes using the expanded chromatin mask depicted in e. **g**, Total number of Ki-67 molecules within the whole DNA volume using the expanded chromatin mask depicted in e. **h**, Density of Ki-67 molecules on the chromosome surface using the outer rim mask depicted in e. Scale bars, $10 \mu\text{m}$.



Extended Data Figure 9 | Dual-colour labelling of the protein termini of Ki-67 indicates an extended conformation oriented perpendicular to the mitotic chromosome surface. **a**, Schematic of Ki-67 labelled at the N terminus (red) and at the chromatin-binding C terminus (green). At high densities, Ki-67 might acquire extended brush-like conformations, as known for polymeric surfactants. **b**, Multispectral fluorescent beads with 500 nm diameter were imaged as reference data to determine dual-colour localization accuracy. The chromatic register shift in the green and red channel was corrected by image registration. **c**, Line profiles (white line) yielded fluorescence profiles, for green and red channels. A sum of two Gaussian functions was fitted (solid line) for each channel and peak-to-peak distances of the corresponding single Gaussian functions (dotted lines) were measured for green ($d(G)$) and red ($d(R)$) channels. **d**, The accuracy of the localization method as illustrated in **b**, **c** was determined based on the absolute difference between green and red peak-to-peak distances divided by 2. Mean (black line) and standard deviation (whiskers) are indicated. **e–h**, Representative example cells of

the quantification in Fig. 4c. **e**, **f**, Live wild-type HeLa cells expressing EGFP–mCherry–Ki-67. Representative example of a sister chromatid pair oriented perpendicular to the imaging plane (**e**). The line profile (white line) yielded a fluorescence profile across one sister chromatid (**f**). A sum of two Gaussian functions was fitted (solid line) for each channel and peak-to-peak distances of the corresponding single Gaussian functions (dotted lines) were measured for green ($d(G)$) and red ($d(R)$), respectively. **g**, **h**, Live wild-type HeLa cell expressing EGFP–Ki-67–mCherry. Representative example of a sister chromatid pair oriented perpendicular to the imaging plane (**g**). The line profile (white line) yielded a fluorescence profile across one sister chromatid (**h**). A sum of two Gaussian functions was fitted (solid line) for each channel and peak-to-peak distances of the corresponding single Gaussian functions (dotted lines) were measured for green ($d(G)$) and red ($d(R)$), respectively. **i**, Model illustrating that an increased grafting density of Ki-67 at the chromosome surface might gradually increase the polymer brush height due to repulsive forces between the polymers. Scale bars, 1 μ m.

Extended Data Table 1 | Proteins ranked by predicted net charge at pH 7

Rank in human proteome	Hugo_ID	ENSEMBL_ID	Charge
1	ASPM	ENSP00000356379	472.565
2	SRRM2	ENSP00000301740	414.8897
3	NEB	ENSP00000484342	254.7429
4	SRRM1	ENSP00000363510	170.867
5	CCDC168	ENSP00000320232	168.6658
6	ZNF729	ENSP00000469582	161.0149
7	ASH1L	ENSP00000357330	148.6947
8	HRNR	ENSP00000357791	148.3841
9	SRRM5	ENSP00000476253	143.9397
10	C2orf16	ENSP00000386190	135.8782
11	KMT2A	ENSP00000432391	134.2102
12	MUC6	ENSP00000487059	134.1865
13	MKI67	ENSP00000357643	133.9347
14	SFI1	ENSP00000383145	130.5004
15	ZNF91	ENSP00000380272	130.3559
16	NKTR	ENSP00000232978	129.8251
17	FLG	ENSP00000357789	118.11
18	ZC3H3	ENSP00000262577	117.9798
19	ZNF99	ENSP00000380293	115.1412
20	SRRM4	ENSP00000267260	113.6421
21	RBBP6	ENSP00000317872	110.2916
22	PRPF4B	ENSP00000433547	109.6788
23	ZNF208	ENSP00000380315	107.9556
24	ZNF721	ENSP00000428878	105.6381
25	SRRM3	ENSP00000480851	104.7896
26	ZNF808	ENSP00000352846	102.4866
27	C1orf167	ENSP00000414909	101.2286
Charge of core histones			
	HIST1H3A	ENSP00000366999	19.9270
	HIST1H2BA	ENSP00000274764	18.22513
	HIST1H4A	ENSP00000244537	17.9833
	HIST1H2AA	ENSP00000297012	15.70769

CAREERS

POST-PHD New graduates' plans vary by discipline **p.315**

GRANT FUNDING Younger and older applicants have similar success rates **p.315**

NATUREJOBS For the latest career listings and advice www.naturejobs.com

SPENCER ROBERTSON/DESIGN PICS/GETTY



Expat life can give scientists a chance to experience other cultures and make new connections.

GLOBAL JOBS

A taste for travel

Moves in science are common, exhilarating — and challenging. Read on for survival tips.

BY EMILY SOHN

For his work as a palaeoclimatologist, Dominik Fleitmann has moved across oceans and continents 5 times in the past 20 years. Born in Germany, where he earned an undergraduate and a master's degree, he has studied and worked in Switzerland, California, Massachusetts, Switzerland again — and now the United Kingdom for a faculty position at the University of Reading.

With each move, Fleitmann faced discomfiting challenges. He once had to scrape money out of his daughter's piggy bank to pay bills

before an international bank transfer arrived. Another time, his family spent several weeks in one small room of their new home while the rest of the house was being renovated. Many times, in the middle of a relocation, he felt on the verge of a nervous breakdown.

But he has come to recognize the useful lessons that moving has taught him, including the need to remain flexible and resilient — both in the lab and at home. “Every scientist has to deal with unexpected or difficult situations: rejections of proposals or manuscripts, analytical and administrative problems,” he says. “Moving around and adapting to new

circumstances can help train your mental and emotional strength. What doesn't kill you makes you stronger.”

Moving to another lab or country is standard for early-career scientists. They are likely to relocate for graduate school, for one or more postdoc positions and then for a permanent or contract job. By changing workplaces, scientists can certainly reap benefits — new collaborators, cultural insights and stronger problem-solving skills — but leaving one way of life to start another is often stressful. Whether you're slogging halfway around the world or just to another city, you will need to cultivate patience, ►

► open-mindedness and a spirit of adventure. Many junior scientists think that they must travel to advance their career. They worry that advisers, lab heads and potential employers will view them as lazy or unmotivated if they stay put. “You hear throw-away comments like, ‘Ah, that’ll look good on your CV because you moved,’” says Pia Lentini, an ecologist at the University of Melbourne, Australia. She says that Australian ecologists face particularly intense pressure from colleagues to relocate and learn about ecosystems elsewhere that are more common than Australia’s unique island ecology.

But such pressure is not just a Southern Hemisphere thing. Some senior researchers might prioritize outside candidates or outright refuse to hire someone from their own organization. Michael Merrifield, an astronomer at the University of Nottingham, UK, often sits on hiring committees for universities and national fellowship panels. He says that senior faculty members with such objections usually argue that applicants with a history of moving have better shown their ambition and devotion to pursuing the best opportunities in science.

GLOBAL GUMPTION

Unquestionably, there are upsides to moving, especially to another country. Fleitmann found that exposure to unfamiliar cultural norms helped him to expand his intellectual limits. In Germany and Switzerland, he says, he had internalized a risk-averse and perfectionist cultural mentality that views failure as a sign of personal defeat. But after time in the United States, he gained the courage to take risks and try new endeavours, such as learning to use a high-resolution ion microprobe to measure trace elements in long cores of coral.

After he moved back to Europe, that experience led to an international collaboration and co-authorship on a paper published in 2007

that used ancient coral samples to understand the history of soil erosion in East Africa. It remains one of the most exciting projects of his career (D. Fleitmann *et al. Geophys. Res. Lett.* **34**, L04401; 2007). His US experience also helped him to break down his own hierarchical assumptions and become open to learning from fellow students — not just from professors, as he had long thought was the only proper option.

Moving can help researchers to refine the way that they do science, Merrifield says. When he arrived at Nottingham, he noticed that students were writing astronomy papers in passive voice: “The experiment was done.” It struck him as strange because he had grown used to seeing active construction when he was doing his PhD and postdoc in the United States and Canada. He helped to create a style guide for students at his new university, which explains that either tense is acceptable. “If you’ve spent too long looking at one system, you would assume that’s the way it has to be because that’s the way it is,” he says. “You never question it.”

Long-distance moves can provide enduring social and research-related benefits, says Pablo Manavella, a molecular biologist and group leader at the Agrobiotechnology Institute of Santa Fe in Argentina. He met an international group of colleagues during a postdoc in Germany, and stayed in touch with many of them after he returned to Argentina. Those connections have since led to collaborations.

Even more invaluable, they have yielded a generous network of trustworthy friends who offer discussion, consultation and experimental troubleshooting on a daily basis. He brought along techniques that he learned in Europe,

“You learn a lot from talking with people who have a point of view that is totally different from yours.”

including more-efficient ways to grow plants — by using sticks and paper bags to avoid contamination — and to organize equipment (by colour-coding tubes). “You learn a lot from being in the world and talking with people about science who have a point of view that is totally different from yours,” he says. “You can read papers and reach people by mail, but it’s not the same.”

Tight bonds can form quickly between uprooted colleagues, adds Danelle Seymour, now a postdoc at the University of California, Irvine. Not long after she moved from California to study plant genetics for a PhD at the Max Planck Institute for Developmental Biology in Tübingen, Germany, her boyfriend severely cut his hand. A colleague whom they had known for just a few months rode in the ambulance to the hospital, an hour away, to help them to communicate in German. “I think you make friends faster than you would in your own country, and they’re there for you,” she says. She remains close with the colleague today.

ISOLATION BLUES

But trying to settle into a new environment is not always an upbeat, exciting challenge. Sometimes, moving back to one’s home nation or region can bring its own breed of cultural whiplash. When Manavella returned to Argentina with his wife and then-1-year-old son, who was born in Europe, he discovered a medical divide. “The paediatrician in Germany told us, ‘You have to feed this guy more — he’s too skinny,’” Manavella says. “We got to Argentina and the paediatrician said, ‘You’ve got to stop feeding him so much. He’s too fat.’ There were totally different perspectives we had to get used to.”

Stumbling on cultural differences that arise from a change of city can also produce painful epiphanies, even in the same country. When medical student Lisa Strate arrived in Boston, Massachusetts, for a gastroenterology residency and fellowship, she hopped onto the city bus to explore residential neighbourhoods for housing options. On one ride, she noticed another passenger holding a book that she had recently read. She tried to strike up a friendly chat about it — an exchange that would have been commonplace in California’s Bay Area, where she had done her undergraduate and medical-school studies. “She just glared at me and didn’t say anything. I thought, ‘Wow, that would’ve been a good conversation starter. I guess no one here wants to chat on the bus.’ I felt a similar attitude in the grocery store and other situations. That was really hard for me to get used to.”

Even without snubbing, isolation is a common by-product of moving, and it can carry a heavy emotional toll. Strate, now at the University of Washington in Seattle, tries to make friends through her job and children, but admits that it’s slow going. Manavella’s mother died while he was in Germany, and he wasn’t able to get back to Argentina for her funeral. With no



Dominik Fleitmann prepares for a move to the United Kingdom, one of many relocations for his career.

GLOBAL CITIZENRY

Expat advice

- When you're caught up in the scrum and chaos of your move, keep your long-term goals in mind.
- Accept that it will take six months or longer before you start to feel settled in a new place.
- Seek out international-student houses and other resources that support incoming students and researchers at your new institution.
- If you move to another country, learn the language.
- Travel around your area, talk to strangers and try foods that are unfamiliar to you.
- Accept every social invitation that you receive from colleagues and co-workers, no matter how shy or tired you are — especially in the beginning. You can skip gatherings later once you have a broad social network. *E.S.*

family nearby, he was left to mourn on his own and with friends, and he threw himself into work until he could return home at Christmas to be with family and honour her.

To combat isolation at a new destination, experienced nomads suggest accepting every social invitation (see 'Expat advice'). Seymour was not much of a cyclist before she moved to Germany, but she joined long-distance rides of up to 50 kilometres simply because she was asked. It is important not to neglect your social life, even while putting in the extra time that it takes to get started at a new job, so that you can connect with others in the community and learn your way around. "When you've moved somewhere and you don't know anyone, it's easy to work around the clock," she says. "That's not the best way to be in the end."

Finding room-mates who work outside of science can also help you to expand your social network, suggests Merrifield. When he moved to Toronto for his first postdoc, he used a local real-estate agent to find an apartment in a converted house with an eclectic mix of neighbours in the building, including one who worked in finance and several musicians. "That helped broaden my circle of friends beyond the people in my department," he says. "We had the coolest parties."

PLANTING ROOTS

Another downside: the time that it takes to adjust to unfamiliar surroundings can negatively affect work performance, says physicist Philip Moriarty, who completed his undergraduate degree and PhD at Dublin City University before moving for a postdoc to Nottingham, where he is now. As a principal

investigator, Moriarty values postdocs who stay around long enough to give his research group a sense of continuity. From a stable position, he adds, scientists can still expand their horizons with extended lab visits, workshops and conferences.

"It's not a particularly conducive environment for a postdoc to do their best work if they're worrying all the time about where the next position will be and having to drag their families across the world," he says. "If you stay in the same place, you can build on your previous work and establish yourself as a truly independent researcher in a more sustained way than you can if you're moving from post to post."

Indeed, some find that the right decision may well be to stay put. While Strate was working on her research fellowship in Boston, her physician-scientist husband accepted a job offer in Seattle. She was working to launch her career, had research funding and wasn't ready to leave. So for several years, the couple lived 4,000 kilometres apart from each other, even as they cared for their two young children.

Once Strate felt that she had a foot in the door to an established research career, and that she could carry on her work in Seattle, she felt comfortable joining her husband and reuniting her family, while continuing to collaborate with colleagues in Boston. The separation was hard, she says, but she thinks that it was the right decision. "You don't want to be regretful," she says. "It can lead to resentment."

Given all the pros, cons and unknowns, it can be hard to make decisions about if, when and where to move, says Henry Sauermann, an economist who specializes in studies of science and innovation and relocated from Germany to North Carolina before landing his current position at the Georgia Institute of Technology in Atlanta. Before each of his moves, he listed and ranked his priorities, both professional and personal. He suggests considering factors such as lab sociability, access to data, tenure guidelines, proximity to friends and family, and gut feelings.

At the end of the day, there is no escaping the fact that major life decisions are always going to be hard. "I do remember vividly, the night before I had to make the decisions, I felt like I had no idea," Fleitmann says. "But it still came together somehow."

It helps to remember that eventually, there will be a new normal, impossible though it might seem. "You have these difficult moments when you are sitting at your desk, and your brain is so full of the things you still have to do," Sauermann says. At times like this, he found inspiration in a famous Winston Churchill quote: "If you're going through hell, keep going." ■

Emily Sohn is a freelance journalist in Minneapolis, Minnesota.

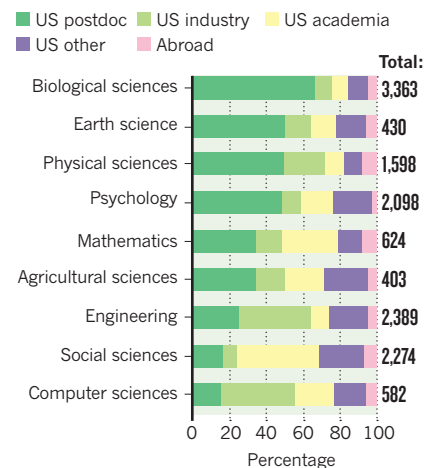
POSTDOCTORAL PLANS

PhD — now what?

Two-thirds of US citizens and permanent residents looking towards their first post after a biology PhD in 2014 said that they planned to do postdoctoral research — the highest proportion in any scientific discipline, a report from the US National Science Foundation has found (see 'Next step on the ladder'). The survey, 'Women, Minorities and Persons with Disabilities in Science and Engineering', also found that biology PhDs outnumbered those in other scientific disciplines. Overall, 42% of science and engineering graduates who had plans said that they intended to do a postdoc, down from 44% in 2012, although the number of PhD recipients rose slightly. In 2001, 39% of graduates with definite plans were heading to a postdoc.

NEXT STEP ON THE LADDER

Overall, 42% of new US PhDs in science and engineering accept postdocs, but specialties vary.



GRANT AWARDS

Age is no advantage

Contrary to common perceptions, the US National Institutes of Health (NIH) awards grants to younger applicants at about the same rate as it does to older ones (M. L. Heggeness *et al. Cell Stem Cell* **19**, 15–18; 2016). Older researchers do win more grants, the study found, but that is because more applicants are older. This could be because scientists receive tenure-track positions later in life, or because younger scientists are taking jobs outside academia, the authors say. For typical 'R01' type grants, applications from people under 40 have declined since the 1980s, but the number of applicants aged 60–64 grew by 40% between 2005 and 2014. In 2014, funding rates were 23% for this group, compared with 25% for researchers aged 35–40 and 19% for researchers under 35.

GLOBAL CITIZENRY

Expat advice

- When you're caught up in the scrum and chaos of your move, keep your long-term goals in mind.
- Accept that it will take six months or longer before you start to feel settled in a new place.
- Seek out international-student houses and other resources that support incoming students and researchers at your new institution.
- If you move to another country, learn the language.
- Travel around your area, talk to strangers and try foods that are unfamiliar to you.
- Accept every social invitation that you receive from colleagues and co-workers, no matter how shy or tired you are — especially in the beginning. You can skip gatherings later once you have a broad social network. *E.S.*

family nearby, he was left to mourn on his own and with friends, and he threw himself into work until he could return home at Christmas to be with family and honour her.

To combat isolation at a new destination, experienced nomads suggest accepting every social invitation (see 'Expat advice'). Seymour was not much of a cyclist before she moved to Germany, but she joined long-distance rides of up to 50 kilometres simply because she was asked. It is important not to neglect your social life, even while putting in the extra time that it takes to get started at a new job, so that you can connect with others in the community and learn your way around. "When you've moved somewhere and you don't know anyone, it's easy to work around the clock," she says. "That's not the best way to be in the end."

Finding room-mates who work outside of science can also help you to expand your social network, suggests Merrifield. When he moved to Toronto for his first postdoc, he used a local real-estate agent to find an apartment in a converted house with an eclectic mix of neighbours in the building, including one who worked in finance and several musicians. "That helped broaden my circle of friends beyond the people in my department," he says. "We had the coolest parties."

PLANTING ROOTS

Another downside: the time that it takes to adjust to unfamiliar surroundings can negatively affect work performance, says physicist Philip Moriarty, who completed his undergraduate degree and PhD at Dublin City University before moving for a postdoc to Nottingham, where he is now. As a principal

investigator, Moriarty values postdocs who stay around long enough to give his research group a sense of continuity. From a stable position, he adds, scientists can still expand their horizons with extended lab visits, workshops and conferences.

"It's not a particularly conducive environment for a postdoc to do their best work if they're worrying all the time about where the next position will be and having to drag their families across the world," he says. "If you stay in the same place, you can build on your previous work and establish yourself as a truly independent researcher in a more sustained way than you can if you're moving from post to post."

Indeed, some find that the right decision may well be to stay put. While Strate was working on her research fellowship in Boston, her physician-scientist husband accepted a job offer in Seattle. She was working to launch her career, had research funding and wasn't ready to leave. So for several years, the couple lived 4,000 kilometres apart from each other, even as they cared for their two young children.

Once Strate felt that she had a foot in the door to an established research career, and that she could carry on her work in Seattle, she felt comfortable joining her husband and reuniting her family, while continuing to collaborate with colleagues in Boston. The separation was hard, she says, but she thinks that it was the right decision. "You don't want to be regretful," she says. "It can lead to resentment."

Given all the pros, cons and unknowns, it can be hard to make decisions about if, when and where to move, says Henry Sauermann, an economist who specializes in studies of science and innovation and relocated from Germany to North Carolina before landing his current position at the Georgia Institute of Technology in Atlanta. Before each of his moves, he listed and ranked his priorities, both professional and personal. He suggests considering factors such as lab sociability, access to data, tenure guidelines, proximity to friends and family, and gut feelings.

At the end of the day, there is no escaping the fact that major life decisions are always going to be hard. "I do remember vividly, the night before I had to make the decisions, I felt like I had no idea," Fleitmann says. "But it still came together somehow."

It helps to remember that eventually, there will be a new normal, impossible though it might seem. "You have these difficult moments when you are sitting at your desk, and your brain is so full of the things you still have to do," Sauermann says. At times like this, he found inspiration in a famous Winston Churchill quote: "If you're going through hell, keep going." ■

Emily Sohn is a freelance journalist in Minneapolis, Minnesota.

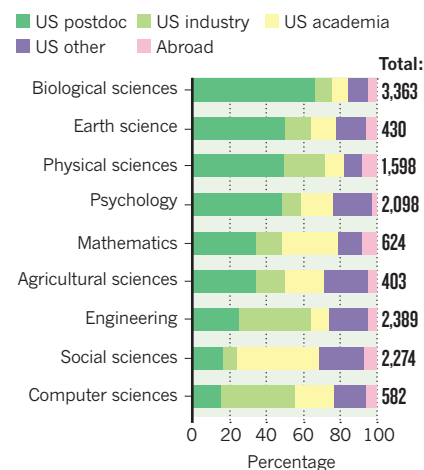
POSTDOCTORAL PLANS

PhD — now what?

Two-thirds of US citizens and permanent residents looking towards their first post after a biology PhD in 2014 said that they planned to do postdoctoral research — the highest proportion in any scientific discipline, a report from the US National Science Foundation has found (see 'Next step on the ladder'). The survey, 'Women, Minorities and Persons with Disabilities in Science and Engineering', also found that biology PhDs outnumbered those in other scientific disciplines. Overall, 42% of science and engineering graduates who had plans said that they intended to do a postdoc, down from 44% in 2012, although the number of PhD recipients rose slightly. In 2001, 39% of graduates with definite plans were heading to a postdoc.

NEXT STEP ON THE LADDER

Overall, 42% of new US PhDs in science and engineering accept postdocs, but specialties vary.



GRANT AWARDS

Age is no advantage

Contrary to common perceptions, the US National Institutes of Health (NIH) awards grants to younger applicants at about the same rate as it does to older ones (M. L. Heggeness *et al. Cell Stem Cell* **19**, 15–18; 2016). Older researchers do win more grants, the study found, but that is because more applicants are older. This could be because scientists receive tenure-track positions later in life, or because younger scientists are taking jobs outside academia, the authors say. For typical 'R01' type grants, applications from people under 40 have declined since the 1980s, but the number of applicants aged 60–64 grew by 40% between 2005 and 2014. In 2014, funding rates were 23% for this group, compared with 25% for researchers aged 35–40 and 19% for researchers under 35.

THE DEPARTMENT OF CORRECTION

A lesson learned.

BY NINAN TAN

“Mr Smith? Your safe room is ready.” The nurse motions me to follow her.

I toss the magazine back on the waiting-room pile, push up from the chair and grab my briefcase. The nurse indulges me with a courtesy smile as I follow her down the door-lined hallway. A man’s muffled cries from behind the padded walls startle me.

“Is this your first time with us?” the nurse asks.

“Yes. Is there anything I should know beforehand?”

“Don’t worry, Mr Smith. Our professional staff will make sure that your session goes nice and smoothly.”

As we reach the door to my room, a scuffle at the other end of the hallway grabs my attention. My heart knocks in my throat as a young woman, bawling hysterically, is removed from one of the safe rooms.

“Repeat offender. Multiple sentences,” the nurse remarks as she unlocks my door.

“Does it hurt?”

“It really depends on the crime,” she says, ushering me into the white, padded room. “May I suggest you remove your suit jacket, Mr Smith?”

I slip my arms out of the jacket and hand it to her. She folds it over her arm and waits until we are joined in the room by an Exec. He barely looks me in the eyes as he fixes a headset to my scalp. My stomach twists with anxiety.

“Let me know if it’s too tight,” he says as he adjusts the headband.

I glimpse at the name tag pinned to his chest. “It’s fine. Thank you, Eric.”

The Exec’s eyes shift to meet mine in surprise. I doubt many notice his name.

“Alright, Mr Smith,” the Exec says. “Once we leave the room and shut the door, there will be a 30-second countdown and then your session will begin.”

“I understand.”

The nurse and the Exec depart and the door locks shut behind them, triggering the longest 30 seconds of my life.

At the end of the countdown, the headset comes to life and I’m no longer in the room. Instead, I find



myself in a Brooklyn street. It’s night, and the pavement is dotted with puddles from an early afternoon shower. I hear the deafening wail of an ambulance as it rushes by. As I inhale, the smell of damp cement and wet trash hits the back of my throat. I shiver from a gust of wind and pop my collar up.

I turn the corner and climb the steps up to a pre-war brownstone building, fumbling for my keys. Inside, the house is newly renovated with a modern touch. A blonde woman in a long, blue dress glides down the stairs. A wave of adoration warms my chest at the sight of her.

“Robert. You’re back. I thought you’d be late again tonight,” she says.

“I couldn’t. Not on our anniversary,” I say. But it’s not my voice. It’s Robert’s. I pull out a bouquet of flowers and offer it to her.

“Oh, Robert, darling.” She takes the flowers from me and presses her soft lips to mine.

The bell rings.

“Who could it be at this time?” she asks.

I open the door and my chest tightens from a mixture of surprise and disquiet. Standing at the top of the steps is me — Simon Smith.

“Mr Smith? What are you doing at my house? I thought we were done talking.”

Simon Smith doesn’t utter a word. The raw redness in his troubled eyes is evident even in his partially shaded face. His hands hide inside his overcoat pockets, tense and curled into fists.

I swallow, uneasy.

“Robert, who is it?” the woman asks from behind me.

I turn to her. “It’s alright, Vanessa. I’ll take care of it.”

Her eyes open wide in shock as a disorientating pain rips through my abdomen. I look down at the knife lodged in my stomach. Blood instantly saturates my white shirt. I topple to the floor. Vanessa’s screams send a rippling wave of panic through me. Fear, confusion and hatred overwhelm my senses as my eyes fix on Simon Smith, my murderer. The pain intensifies. Jagged white lights cut across my vision, until everything fades to black.

I come to, gasping for breath. The lock clicks, and the Exec and the nurse re-enter the safe room. The Exec kneels beside me and removes the headset. My heart continues to pound and I hug myself, trembling.

“Your first session is complete, Mr Smith,” the Exec says. “Elapsed time is 3 minutes and 17 seconds.” He grabs my arms, hoisting me up from the floor. My shirt is soaked in sweat. “The first time always proves to be the most demanding. But don’t worry. Your nervous system will gradually adjust to the program.”

The nurse grabs my briefcase and loops her arm around mine as she helps me out of the safe room and back to the front desk.

The receptionist hands me a receipt for my session. “Mr Smith,” she says. “As required by law, after every session, I am obligated to remind you of why you are here: you have been convicted of the murder of one Robert Clement on 10 October 2051. Your sentence is to relive the aforementioned murder from the perspective of the victim, once a week, every week, for 25 years. You currently have 1,299 sessions left.”

Behind me, the nurse still smiles. “Thank you for choosing Clarion Correctional Facilities, Mr Smith. We hope to see you again next week. Please, don’t forget to fill out our customer satisfaction card for a chance to win an upgrade to our deluxe suites. Have a good rest of your day.” ■

Ninan Tan is a social-issues filmmaker and science-fiction writer based in New York City. She also frequently migrates to Rome and Beijing in the summers and winters. Learn more at: www.theorphanclassifieds.com.

ILLUSTRATION BY JACEY